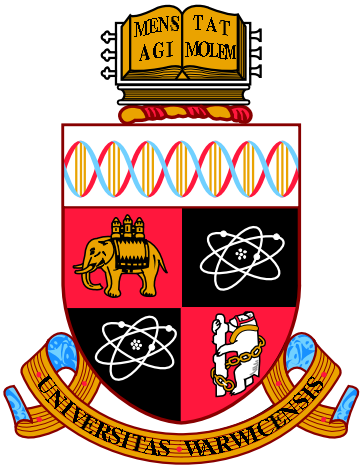


Search for high-mass dimuon resonances produced in association with b-quarks at the ATLAS detector

by

Christopher John McNicol

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Physics

July 2019

THE UNIVERSITY OF WARWICK

Contents

List of Tables

List of Figures

Acknowledgments

There are many people I would like to thank for their support and friendship throughout these challenging years. Firstly, I would like express my gratitude to Prof. Sinéad Farrington for supervising and guiding me. I would also like to thank Profs. Bill Murray and Paul Harrison for their useful advice. I owe a great deal of appreciation to Dr. Elisabetta Pianori and Dr. Ligang Xia, for their help and most importantly, their patience. Furthermore I am particularly grateful to Dr. Tim Martin for his willingness to answer my many questions. I must also thank my friends in the office: Adomas, Martin, Van, Bryn, Eleanor, Cayo, Ed and Andy, for many hours of interesting conversations which were only sometimes useful but were always entertaining.

I also want to express my deepest thanks to all of my family and friends. My long-suffering wife has supported me through the best and worst of times and I love her very much for that. I also wish to thank my mother for her continued kind encouragement and willingness to listen. Finally I would like to thank my father, whose curiosity and passion for learning I have inherited, I hope that this thesis is worthy to enter his vast library. This is for you, Dad.

Declarations

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

My contribution to the online running of ATLAS and qualification task to become an author of the collaboration was the development and maintenance of the ATLAS High-Level Trigger (HLT) and in particular the Automated Cost Monitoring (AutoCostMon) described in Section 2.5 .

The search for scalar resonances decaying into dimuon pairs and produced in association with b-quarks [1], which forms a large part of this thesis, was a collaborative effort involving five ATLAS members from different institutes. As the main analyser, my contributions include:

- development of the analysis framework in C++ and python using ROOT [2] and RooFit [3]
- production of simulated samples of signal events (Section 4.2)
- definition and optimisation of signal and control regions (Section 5.2.2)
- implementation of systematic uncertainties provided by ATLAS Combined Performance groups (Section 7.1)
- estimation of analysis-specific systematic uncertainties (Section 7.2)
- development of the signal interpolation procedure (Section 6.3)

A paper describing this analysis was published in the Journal for High Energy Physics [1].

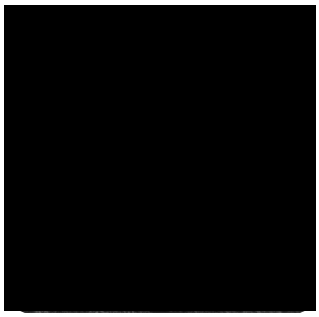
Abstract

A search for heavy dimuon resonances produced in association with b-jets is presented. The analysis uses a 36.1 fb^{-1} dataset collected in 2015 and 2016 by the ATLAS detector during Run 2 of the LHC at $\sqrt{s} = 13$ TeV. No significant excess in events above the expected background from Standard Model processes was observed. Upper limits on the cross-section times branching ratio were set at 95% Confidence Level for two production mechanisms: gluon-gluon fusion and b-associated production. The observed upper limits are between 1.6 and 44 fb and 1.9 and 41 fb respectively and agree well with expectations.

Introduction

If we knew what it was we were doing, it would not be called research, would it?

Albert Einstein



The discovery of the Higgs Boson in 2012 [4,5] is arguably the crowning achievement of the Standard Model of particle physics and the result of decades of developments in understanding the fundamental constituents and interactions of the Universe. The Standard Model successfully describes three out of four fundamental forces of nature and classifies all known elementary particles. However, there are still many open questions to be answered, not the least of which is whether this observed boson is part of an extended Higgs sector. There are many reasons to believe that the Standard Model is not complete in its current state and so there is significant effort in searching for experimental evidence for proposed extensions. Even null results are useful as they eliminate possible models and allow for refined and better targeted searches in the future.

This thesis describes just such a search, in particular looking for resonances decaying into a muon-antimuon pair with and without associated b-jets. Rather than searching for a particular extension to the Standard Model, from its conception this search was designed to be *model-independent*. If no significant deviation from Standard Model expectations is observed, the objective is to find model-independent

limits on the production cross-section that can be used to exclude many different models. The data analysed herein was collected by the ATLAS experiment in 2015 & 2016 during Run 2 of the Large Hadron Collider (LHC) and comprises an integrated luminosity of 36.1 fb^{-1} of proton-proton collisions. The centre-of-mass energy of the Large Hadron Collider (LHC) of $\sqrt{s} = 13$ TeV is the highest achieved by any particle collider to date, and allows unprecedented sensitivity to physics at the electroweak scale.

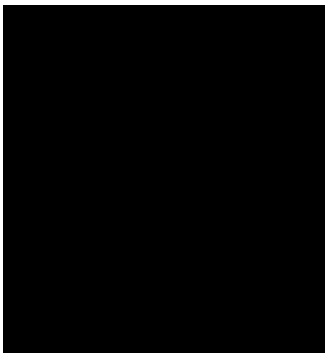
This thesis shall begin with an overview and explanation of the theoretical underpinnings of modern particle physics in Chapter 1, with an additional summary of possible extensions and motivations for the search. This is followed in Chapter 2 by a description of the experimental facilities used, namely the CERN accelerator complex including the Large Hadron Collider and the ATLAS experiment. The algorithms used to identify and reconstruct particles in the detector are explained in Chapter 3 along with how measurements of these particles are calibrated. In Chapter 4 the simulation of signal and background processes - used to guide the analysis - is described. Chapters 5 & 6 respectively discuss details on the selection criteria and the statistical techniques used in this analysis. The systematic uncertainties considered are outlined in Chapter 7 and finally the results of the search are discussed in Chapter 8.

Chapter 1

Theory

In theory there is no difference
between theory and practice.
But in practice there is...

Yogi Berra



1.1 The Standard Model

The Standard Model (SM) is a description of the fundamental particles of nature and how they interact. It is the result of decades of theoretical refinement and experimental investigation, starting with the early pioneers of Quantum Mechanics and resulting in the recent discovery of the Higgs boson.

One of the first steps was to develop a new theoretical framework that was consistent with both Quantum Mechanics and Special Relativity, two of the most successful theories of the early 20th Century. Quantum Mechanics described how physics at the scales of molecules, atoms and even smaller is very different from the classical, macroscopic world of our intuition. Through special relativity, Albert Einstein derived how the measurements made by observers in different frames of

reference can differ. The result of work by Dirac among many others, to unify these theories was the development of Quantum Field Theory (QFT).

This section will start with a brief overview of the particle content of the Standard Model, including their properties. Then a summary of QFT will be given, defining concepts and terminology which will be useful for discussing the Standard Model in more detail. The descriptions of the electromagnetic, weak and strong forces using QFT will be briefly explained, in particular focusing on how these interactions can arise from fundamental symmetries of nature. The final part of the Standard Model that will be discussed is the Higgs mechanism. At the end of this chapter, the problems and shortcomings of the Standard Model will be described along with possible extensions. Finally, a motivation for the experimental search that forms a large part of this thesis shall be given.

1.1.1 Standard Model Particles

The fundamental particles of nature have different quantum numbers that dictate their behaviour, and they can be separated into several groups as shown in Figure 1.1 and Table 1.1. The Stern-Gerlach experiment [6] demonstrated that particles have intrinsic angular momentum, known as *spin*, and that it is quantised in units of $\hbar$. Fermions are particles that have half-integer spin whereas bosons have integer spin, each displaying different properties via the spin-statistics theorem [7]. The 'matter' particles of the Standard Model are fermions and the 'force' particles are bosons. The fermions can be further divided by the types of interactions they can be involved in; quarks interact via the strong interactions whereas leptons do not. The interactions a particle can undergo are dictated by their quantum numbers which in this context are called 'charges', in analogy with the electric charge in electromagnetism. There are three generations of fermions, which only differ in their masses. As of yet, there is no explanation for why there should be three generations, rather than two or four, for instance. The other major group of particles are the gauge bosons which are mediators for the strong, weak and electromagnetic interactions. One of the greatest successes of high-energy physics in the 20th century was to identify the link between the fundamental forces and symmetries of the Universe. Before discussing this further, it will be useful to give a brief description of the different types of fundamental particles in the Standard Model.

Figure 1.1: The fundamental particles of the Standard Model, arranged into quarks, leptons, gauge bosons and the Higgs boson.

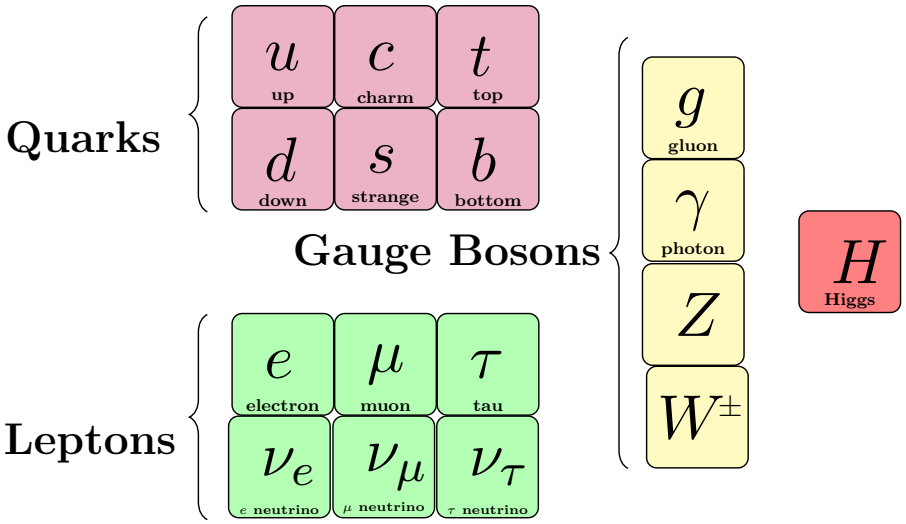


Table 1.1: Properties of the fundamental particles [8].

Particle	Symbol	Spin	EM charge / e	Weak Isospin	Colour Charge	Mass /MeV
electron	e	1/2	-1	$-1/2$	0	0.511
electron neutrino	ν_e	1/2	0	$+1/2$	0	$< 2 \times 10^{-6}$
muon	μ	1/2	-1	$-1/2$	0	105.7
muon neutrino	ν_μ	1/2	0	$+1/2$	0	< 0.19
tau	τ	1/2	-1	$-1/2$	0	1.777×10^3
tau neutrino	ν_τ	1/2	0	$+1/2$	0	< 18.2
up	u	1/2	$+2/3$	$+1/2$	r,g,b	≈ 2.3
down	d	1/2	$-1/3$	$-1/2$	r,g,b	≈ 4.8
charm	c	1/2	$+2/3$	$+1/2$	r,g,b	$\approx 1.275\times^3$
strange	s	1/2	$-1/3$	$-1/2$	r,g,b	≈ 95
top	t	1/2	$+2/3$	$+1/2$	r,g,b	$\approx 173 \times 10^3$
bottom	b	1/2	$-1/3$	$-1/2$	r,g,b	$\approx 4.2 \times 10^3$
gluon	g	1	0	0	Colour-anticolour pair	0
photon	γ	1	0	0	0	0
Z boson	Z^0	1	0	0	0	91.188×10^3
W boson	$W^\pm$	1	± 1	± 1	0	80.4×10^3
Higgs boson	H^0	0	0	$+1/2$	0	125.1×10^3

Leptons

The electron was the first subatomic particle discovered when J.J. Thomson measured the mass of cathode rays in 1897 and it later became an invaluable tool in developing Quantum Mechanics [9]. Its properties, including electric charge and mass were accurately measured and its interactions with other matter were studied. One interaction of particular interest to Wolfgang Pauli in 1930 was beta decay - the radioactive decay of a nucleus leading to the production of an electron. Energy

and momentum did not appear to be conserved and so Pauli proposed the *neutrino*, an undetected particle which possessed the missing energy. Then in 1936, Carl Anderson and Seth Neddermeyer discovered a particle that looked much like an electron but with a significantly larger mass [10]. These particles were eventually grouped into the *leptons* - particles that do not interact via the strong nuclear force - and since then the tau lepton and two more generations of neutrino have been discovered [11–13].

The Standard Model describes three generations of leptons, each comprising a pair of a charged lepton and a neutrino which can interact with each other via the weak interaction. This coupling is dependent on the chirality of the particle (see Section 1.1.3). Right-handed fermions and left-handed antifermions are weak-isospin singlets and do not undergo weak interactions whereas left-handed fermions and right-handed antifermions form weak-isospin doublets and do therefore interact.

The couplings of the leptons to the gauge bosons has long believed to be independent of their generation - a concept known as *lepton universality*. However, some recent tests of lepton universality in the decays of B mesons at Belle, BaBar and LHCb have shown deviations from the Standard Model, whilst others do not show any discrepancy [14–21]. There is still some controversy about these results and how they should be collectively interpreted . Another recent revelation is that neutrinos are not massless. Instead, the small mass difference between neutrino flavours leads to *neutrino oscillations*, the first observations of which were awarded the 2015 Nobel Prize in Physics [22].

Quarks

Quarks are the only fermions with $SU(3)$ 'colour' charge and therefore interact via the strong force. The concept of 'colour-confinement' requires that this colour charge is not directly observable at temperatures below the Hagedorn temperature [23] (approximately 3 Trillion Kelvin). Instead, quarks form colour-singlet, bound states called hadrons. These bound states are usually constructed from quark-antiquark pairs, known as mesons, or three quarks which are called baryons. More exotic bound states have been proposed and recent results claim the discovery of so-called tetraquarks and pentaquarks [24–27]. The quark model was first proposed independently by Murray Gell-Mann and George Zweig in 1964 [28–30] and subsequent experimental results showed the proton had substructure [31]. Initially only the three lightest quarks, u, d, s were included but developments in theory and experiments led to the discovery of the three heavier quarks, c, b, t [32–36] with the heaviest 'top' quark discovered in 1995 [37, 38].

Just as for the leptons, left-handed quarks can be grouped into isospin doublets whereas right-handed quarks are isospin singlets. The 'up-type' quarks have electromagnetic charge $+2/3e$ and 'down-type' quarks have charge $-1/3e$ so that every bound state has in total an integer multiple of the elementary charge, e.

Gauge bosons

Bosons are particles with integer spin and each of the fundamental forces in the Standard Model is associated with at least one gauge boson. They are the fundamental quanta of the gauge fields which come about due to symmetries in nature. This link between symmetries and the gauge bosons is an important topic covered in the following sections. The first of these bosons is the photon, the quanta of light. Whilst Descartes and Newton proposed that light was composed of 'corpuscles' in the 17th century, it was not until the 20th century - when Planck and Einstein used the quantisation of light to explain blackbody radiation and the photoelectric effect - that there was strong experimental evidence for this phenomenon. The interaction between electrons and photons was the focus of many developments in the early 20th century culminating in the first successful QFT, Quantum Electrodynamics (QED).

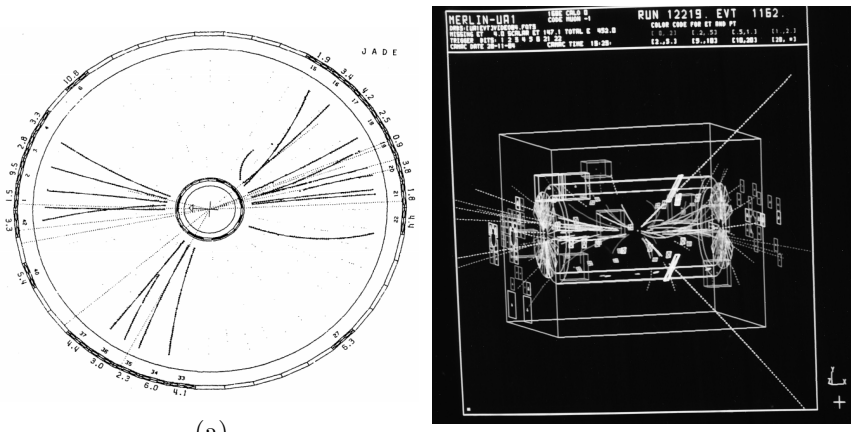
After this success, theorists sought to describe the weak and strong nuclear interactions in the same way. This led to the prediction of the gluon which mediates the strong force and the $W^{\pm}$ and Z^0 bosons which are involved in weak interactions. The gluon was discovered in 1978 in events with 3-jet topologies [39] (see Figure 1.2) with the W and Z bosons discovered shortly after in 1983 [40, 41].

Higgs boson

The final particle in the Standard Model is the Higgs boson, which was only recently discovered in 2012 by the ATLAS and CMS collaborations [4, 5]. The Higgs boson, the quanta of the Higgs field, is not a gauge boson. Unlike the gauge fields, which are spin-1, measurements so far indicate this new boson is spin-0, as predicted [44, 45]. The Higgs field is necessary to describe how the $W^{\pm}$ and Z^0 bosons acquire mass in a process called the Higgs mechanism, or Electroweak Symmetry Breaking (EWSB) which is described in greater detail in Section 1.1.4.

1.1.2 Quantum Field Theory formalism

Before discussing the specifics of the Standard Model, it will be useful to briefly discuss Quantum Field Theory in general. QFTs extend classical field theories by promoting the fields to the role of quantum mechanical operators. In a classical



(a) (b)

Figure 1.2: Experimental evidence for the gluon and Z boson. The 3-jet topology of an event in the JADE detector (a) is indicative of the presence of a gluon-jet [42]. The white tracks on the event display of the UA1 detector (b) are an electron-positron pair produced in the decay of a Z^0 boson [43].

field theory, the equations of motion are determined by minimising the total *action* of the system:

$$S = \int L dt = \int \mathcal{L}(x, \phi, \partial_\mu \phi) d^4 x, \tag{1.1}$$

where L and $\mathcal{L}$ are the *Lagrangian* and *Lagrangian density* respectively. Minimising the action is equivalent to requiring that $\delta S = 0$. If one differentiates the integrand of eqn. (1.1), the results are the Euler-Lagrange equation for classical fields:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = 0. \tag{1.2}$$

In QFTs, the path integral formulation can be used to calculate the scattering amplitude for a given process. To calculate the probability amplitude for a state $|\phi_I\rangle$ at time $t = 0$ to evolve to a final state $|\phi_F\rangle$ at time $t = T$ one divides the total time T into N smaller intervals.

$$\langle \phi_F | \, e^{-iHT} \, | \phi_I \rangle = \int d\phi_1 \cdots \int d\phi_{N-1} \, \langle \phi_F | \, e^{-iHT/N} \, | \phi_{N-1} \rangle \cdots \langle \phi_1 | \, e^{-iHT/N} \, | \phi_I \rangle \,, \tag{1.3}$$

where H is the *Hamiltonian*. In the limit $N \to \infty$ one obtains the Feynman path integral:

$$\langle \phi_F | \, e^{-iHT} \, | \phi_I \rangle = \int \mathcal{D}\phi(t) e^{iS[\phi]} = \int \mathcal{D}\phi(t) e^{\{i \int_0^T dt L\}}, \tag{1.4}$$

which represents an integral over all possible 'paths', $\mathcal{D}\phi(t)$, that the field ϕ can take. In practice, one often uses perturbation theory to perform this calculation; by

expanding the Lagrangian in terms of the interaction couplings, which are assumed to be small, and calculating only the first few terms. The terminology used is that a calculation is performed to Leading Order (LO), Next-to-Leading Order (NLO) etc.

Cross-section calculations

In order to be a useful theory it must be possible to calculate measurable quantities, and in particular, the likelihood of finding a particular final state after a collision. One can do this using the *cross-section*, a quantity that is intrinsic to colliding particles.
Firstly, one can represent a particle as a wavepacket using its fourier transform:

$$|\phi\rangle = \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2E_k}} \phi(\vec{\boldsymbol{k}}) \, |\vec{\boldsymbol{k}}\rangle \, . \tag{1.5}$$

The matrix in the scattering amplitude definition given in eqn. (1.4) can be separated into the uninteresting case where the particles do not interact - where it is just the identity matrix and the part where interactions do occur:

$$\langle \phi_F | \, e^{-iHT} \, |\phi_I\rangle \equiv \langle \phi_F | \, \mathbb{1} + iT \, |\phi_I\rangle \, . \tag{1.6}$$

By applying 4-momentum conservation, one can define the *invariant matrix element*, $\mathcal{M}$:

$$\langle \phi_F | \, iT \, |\phi_I\rangle = (2\pi)^4 \delta^{(4)}(k_A + k_B - \Sigma p_f) \times i\mathcal{M}(k_A, k_B \to p_f), \tag{1.7}$$

where k_A, k_B are the initial state particle momenta and p_f are the final state momenta. The differential cross-section can now be defined:

$$\begin{aligned} d\sigma &= \frac{1}{2E_A 2E_B |v_A - v_B|} \left(\prod_f \frac{d^3p_f}{(2\pi)^3} \frac{1}{2E_f} \right) \times \mathcal{M}(k_A, k_B \to p_f)(2\pi)^4 \delta^{(4)}(k_A + k_B - \Sigma p_f) \\ &\equiv \frac{1}{2E_A 2E_B |v_A - v_B|} \; \mathcal{M}(k_A, k_B \to p_f) \; d\Pi_n, \end{aligned} \tag{1.8}$$

where the cross-section dependencies on the energies and velocities of the initial particles (E_A, E_B, v_A, v_B); the possible final state momenta and the interaction Hamiltonian have been separated. The first term is called the *flux* term and describes the energies and frequency of collisions between particles. The second term, the *matrix element* $\mathcal{M}$, describes the probability of an interaction taking place, and the final term, $d\Pi_n$ is the *phase space* which describes all of the possible final-state momenta that the particles could take.

Representation of Lorentz group

The Lagrangian must be relativistically invariant, that is it must be unchanged by the Lorentz transformation $x^\mu \to \Lambda^\mu_\nu x^\nu$. Since successive Lorentz boosts can be expressed as a single Lorentz transformation, $\Lambda'\Lambda = \Lambda''$, they form a group. It is natural, therefore, to find a matrix representation for this group that can be used when writing terms in the Lagrangian. A basis of sixteen 4×4 matrices can be constructed as seen in Table 1.2, organised by their transformation properties under the Lorentz group.

Table 1.2: Summary of matrix representation of Lorentz group

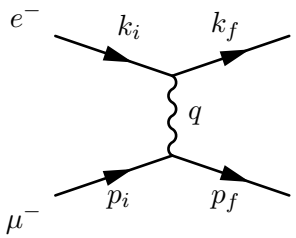
Matrices	Transforms as	Number
$\mathbb{1}_4$	Scalar	1
γ^μ	Vector	4
$\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$	Tensor	6
$\gamma^\mu\gamma^5$	Pseudo-vector	4
γ^5	Pseudo-scalar	1

where, in the 'chiral' basis, γ^μ and γ^5 are defined in 2×2 block form using the Pauli sigma matrices as:

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1}_2 \\ \mathbb{1}_2 & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad \gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3. \tag{1.9}$$

Feynman Diagrams

A helpful way of representing particle interactions is to use a Feynman diagram. Each diagram represents a term in the expansion of the perturbative S-matrix; the lines correspond to particles in *momentum* space - straight lines for fermions and curly for gauge bosons. A set of Feynman rules - which are dependent on the particular QFT theory one considers - can be used to write down the terms needed to calculate the contribution to the matrix element. For instance, the leading order term for electron muon scattering can be represented by the tree diagram:



$$\implies [\overline{u}(p_f)(-ie\gamma^\nu)u(p_i)]\left[\frac{-ig_{\nu\rho}}{q^2}\right][\overline{u}(k_f)(-ie\gamma^\rho)u(k_i)]\,. \tag{1.10}$$

1.1.3 Symmetries and fundamental forces

A symmetry of the Lagrangian is a transformation which leaves the Lagrangian unchanged. Noether's theorems state that every differentiable symmetry of the Lagrangian leads to a conserved quantity [46]. Simple examples are translational and rotational invariance which give rise to the conservation of linear and angular momentum. In the early days of Quantum Field Theory, a number of symmetries were assumed; charge (C), parity (P) and time (T). In fact, it was shown in 1957 by C.S. Wu that parity is *maximally* violated in the decays of cobalt-60 [47]. However, it was still possible that CP - the combined action of the charge and parity operators - would remain as a symmetry. More recently it has been shown that even CP symmetry is broken in the quark sector [48] with hints that it may be violated in neutrinos also [49]. At present it is believed that CPT symmetry must hold as creating a Lorentz-invariant QFT is impossible without maintaining it.

The symmetries that have been discussed so far are all global symmetries. Local symmetries - where transformations are dependent on space-time position - give rise to gauge fields which describe the interactions in the Standard Model.

Quantum Electrodynamics

QED was the first QFT theory used to describe interactions in nature; in particular, the electromagnetic interaction. It is perhaps the simplest example of how a local symmetry of the Lagrangian leads to a gauge field. Starting with the Dirac Lagrangian for a fermion, represented by the *spinor* ψ, in free space [50]:

$$\mathcal{L} = \psi^\dagger \gamma^0 (i\gamma^\mu \partial_\mu - m)\psi \tag{1.11}$$

$$\equiv \overline{\psi}(i\partial\!\!\!/ - m)\psi, \tag{1.12}$$

one can see that the global transformation $\psi \rightarrow e^{-iq\alpha}\psi$ leaves the Lagrangian unchanged, however when requiring a similar *local* symmetry, dependent on space-time coordinates such as

$$\psi \rightarrow e^{-iq\alpha(x)}\psi, \tag{1.13}$$

one introduces a new term to the Lagrangian:

$$\mathcal{L} \rightarrow \mathcal{L} + q\overline{\psi}\gamma^\mu \psi \partial_\mu \alpha. \tag{1.14}$$

By defining a new field, $A_\mu(x)$ which simultaneously transforms as $A_\mu(x) \rightarrow A_\mu(x) + \partial_\mu \alpha$, one can construct an invariant Lagrangian:

$$\mathcal{L} = i\overline{\psi}\partial\!\!\!/\psi - m\overline{\psi}\psi - q\overline{\psi}\gamma^{\mu}\psi A_{\mu} \tag{1.15}$$

$$\equiv \overline{\psi}(i\mathcal{D}\!\!\!/ - m)\psi, \tag{1.16}$$

where the *gauge covariant derivative* $\mathcal{D}_{\mu} \equiv \partial_{\mu} + iqA_{\mu}$ has been used. In the context of electromagnetism, the field A_{μ} represents the photon and the additional term in the Lagrangian represents the interaction of the photon with the fermion. This interaction has a strength governed by the coupling, q. Since $\overline{\psi}\gamma^{\mu}\psi$ is a Lorentz vector (see Table 1.2) one can infer that the introduced photon field is also a vector so that the contraction is a Lorentz scalar. This field is analogous to the four-potential used in classical electromagnetism, and following this analogy one can add the electromagnetic tensor $F_{\mu\nu} \equiv \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$ which governs the propagation of the photon. This results in the complete Lagrangian for QED

$$\mathcal{L}_{QED} = \overline{\psi}(i\mathcal{D}\!\!\!/ - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{1.17}$$

In summary, maintaining the local gauge symmetry of the Lagrangian requires the introduction of a new *gauge field*. The extra terms added to the Lagrangian dictate the interactions and properties of the new field.

Running of couplings and renormalisation

Before going further and studying more complicated QFTs it will be helpful to discuss some important features that can be most easily demonstrated within QED. As previously stated, the matrix element can be calculated perturbatively, adding up terms with successively larger powers of the coupling constants. Higher-order terms contain internal particle loops, where all possible momenta of the loop must be integrated over. Consider the term represented by the Feynman diagram in Figure 1.3. The contribution of this fermion loop to the matrix element is given by:

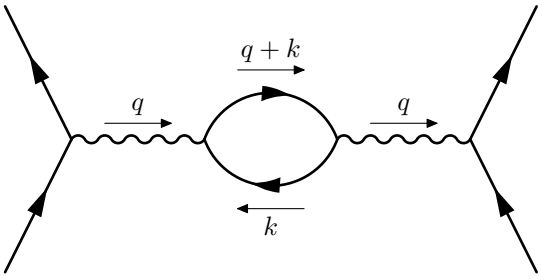


Figure 1.3: Higher-order Feynman diagram containing a particle loop.

$$(-1)(-ie)^2 \int \frac{d^4k}{(2\pi)^4} tr\left[\gamma^\mu \frac{i}{\not{k} - m}\gamma^\nu \frac{i}{\not{k} + \not{q} - m}\right] \tag{1.18}$$

$$\equiv i\Pi_2^{\mu\nu}(q) \equiv i(q^2 g^{\mu\nu} - q^\mu q^\nu)\Pi_2(q^2), \tag{1.19}$$

where in the last term, the functional dependence on q^2 has been separated from the tensor structure[1]. It can be shown [51, p. 246] that these higher-order terms can be included by modifying the tree-level photon propagator like so:

$$\frac{-ig_{\mu\nu}}{q^2} \to \frac{-ig_{\mu\nu}}{q^2(1 - \Pi(q^2))}, \tag{1.20}$$

where $\Pi(q^2)$ is the sum of contributions to infinitely high orders, of which $\Pi_2(q^2)$ is the second-order contribution. For now, it is not important to know what functional form $\Pi(q^2)$ takes, as long as it is regular at $q^2 = 0$. For scattering processes, in a procedure known as *charge renormalisation*, one can reinterpret the effect of the summing the infinite number of corrections by redefining an effective coupling which is q^2-dependent

$$\alpha_{eff}(q^2) = \frac{e^2}{4\pi} \cdot \frac{1}{1 - \Pi(q^2)}. \tag{1.21}$$

The concept that the coupling strength changes or *runs* depending on the energy scale of a scattering process has been experimentally verified [52]; Figure 1.4 shows the effective coupling strength increases at higher momentum scales. A useful intuitive picture is that at low-q^2, virtual e^+e^- pairs act to screen the bare charge but at high-q^2 one can probe closer to the bare charge.

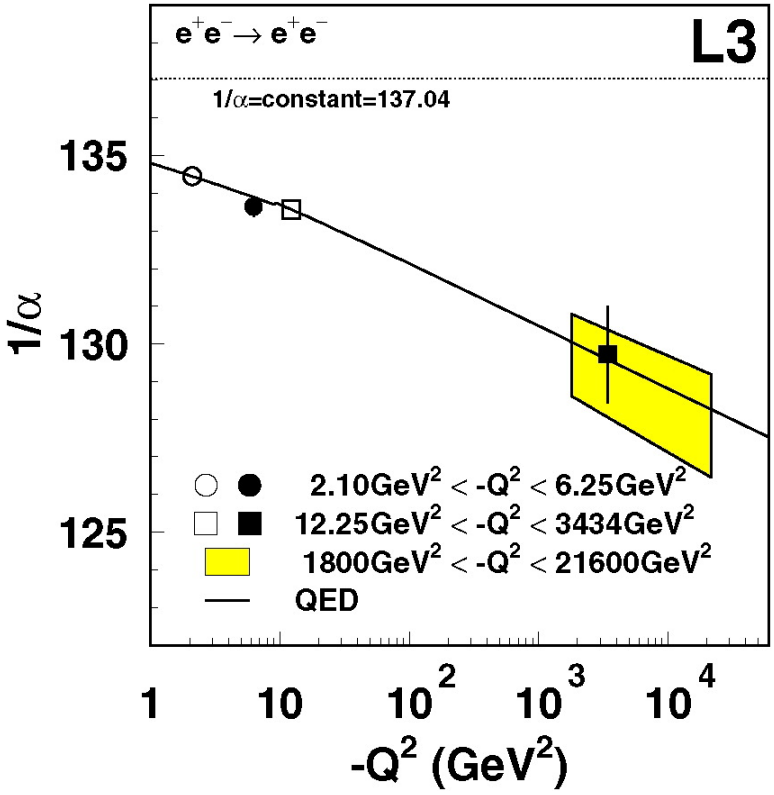
When one calculates the integral $\Pi_2(q^2)$, which is non-trivial, one finds that it has an ultraviolet divergence. This implies that the bare charge of QED is infinite. Whilst philosophically unsettling, the only physically measurable quantity is the effective electric charge and its dependence on q^2 and not the bare charge. Therefore, eqn. (1.21) can be re-written:

$$\alpha_{eff}(q^2) = \frac{e^2}{4\pi} \cdot \frac{1}{1 - \widehat{\Pi}(q^2)} = \frac{e^2}{4\pi} \cdot \frac{1}{1 - [\Pi(q^2) - \Pi(0)]}. \tag{1.22}$$

In this case the divergences of the terms, $\Pi(q^2)$ and $\Pi(0)$ cancel so that the result is finite. In this example, the effective electric charge was shown to have scale-dependent behaviour. In fact, all measurable quantities in a QFT are dependent

[1]This tensor structure can be determined from a *Ward-Takahashi identity*, $q_\mu \Pi^{\mu\nu} = 0$ which is a consequence of the conservation of electric current [51, p. 238–245].

Figure 1.4: Measurement of running of electromagnetic coupling as a function of q^2 at LEP [52]



on the energy scale at which one is measuring them and the exact nature of this dependence can be derived using a renormalisation group. Only in certain theories is it possible to cancel out ultraviolet divergences in the way described above. These theories are called *renormalisable*, an important condition that the theory can provide predictive power. A common belief among high-energy physicists is that the QFTs of the Standard Model represent only 'effective field theories' which are valid up to some large energy scale Λ, wherein a new theory of nature is needed. In this case, a renormalisable theory must be independent of this cut-off scale.

Yang-Mills Theory

In QED the gauge transformation is a member of the $U(1)$ group which is *abelian*, and the transformations are *commutative*. The study of noncommuting local symmetries is called *non-abelian gauge theory* of which Yang-Mills theory [53] is an example that focusses on the $SU(N)$ groups. Yang-Mills theory is important in studying the strong interaction and electroweak unification which are discussed later.

Perhaps the simplest case is for the $SU(2)$ group, where instead of a single fermion field, one considers doublets of Dirac fields and the corresponding Lagrangian:

$$\psi = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \end{pmatrix}, \tag{1.23}$$

$$\mathcal{L} = \overline{\psi}(i\partial\!\!\!/ - M)\psi, \tag{1.24}$$

where M is now a 2×2 matrix. One can see that the field transformation $\psi \rightarrow exp\left(-ig\alpha^i \frac{\sigma^i}{2}\right)\psi$, where σ^i are the Pauli sigma matrices, leaves the Lagrangian unchanged. By requiring that α^i be space-time dependent - in other words enforcing a *local* gauge symmetry - one has to modify the Lagrangian in a similar way to the case of QED, by introducing a new covariant derivative:

$$\mathcal{D}_\mu = \partial_\mu + ig\frac{\sigma^i}{2}A^i_\mu. \tag{1.25}$$

Here three new fields, A^i_μ, have been introduced which - in the limit that α^i is small - must transform according to the following:

$$A^i_\mu \rightarrow A^i_\mu + \partial_\mu\alpha^i + g\epsilon^{ijk}\alpha^j A^k_\mu, \tag{1.26}$$

where ϵ^{ijk} is the *Levi-Cevita symbol* which appears due to the anticommutation of the Pauli sigma matrices. A propagation term can also be constructed with a field strength tensor analogous to the QED case:

$$F^i_{\mu\nu} = \partial_\mu A^i_\nu - \partial_\nu A^i_\mu - g\epsilon^{ijk}A^j_\mu A^k_\nu. \tag{1.27}$$

The final term is again a consequence of the noncommutative nature of $SU(2)$ transformations. Using the definitions of covariant derivative and field strength tensor, the Yang-Mills Lagrangian can be simply written in a very similar manner to QED:

$$\mathcal{L}_{Yang-Mills} = \overline{\psi}(i\not{\mathcal{D}} - M)\psi - \frac{1}{4}F^{i\mu\nu}F^i_{\mu\nu}. \tag{1.28}$$

Quantum Chromodynamics

Whilst the example Yang-Mills Lagrangian above does not describe any interactions in the Standard Model, Quantum Chromodynamics (QCD) - the description of the strong nuclear interaction - is a Yang-Mills theory for the $SU(3)$ group where the generated gauge fields are called *gluons*. In QCD, the conserved quantity is known

as 'colour' and quarks are represented as colour triplets which transform under the $SU(3)$ group [54].

$$\psi = \begin{pmatrix} \psi_r \\ \psi_g \\ \psi_b \end{pmatrix}, \tag{1.29}$$

$$\psi \rightarrow exp\left(-iq\alpha^a \frac{\lambda^a}{2}\right)\psi. \tag{1.30}$$

The labels r, g, b denote an arbitrary colour basis, usually taken as red, green and blue. The λ^a are the Gell-Man matrices which span the Lie algebra of the $SU(3)$ group and obey the commutation relation:

$$\left[\frac{\lambda^a}{2}, \frac{\lambda^b}{2}\right] = if^{abc}\frac{\lambda^c}{2}, \tag{1.31}$$

where f^{abc} are known as the *structure* constants of the group. Imposing local gauge invariance requires the addition of eight new fields, A^a_μ which are the gluons. Following a similar prescription to the Yang-Mills theory described previously, one can write the covariant derivative, field strength tensor and final QCD Lagrangian as

$$\mathcal{D}_\mu = \partial_\mu + ig\frac{\lambda^a}{2}A^a_\mu, \tag{1.32}$$

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - gf^{abc}A^b_\mu A^c_\nu, \tag{1.33}$$

$$\mathcal{L}_{QCD} = \overline{\psi}(i\not{\mathcal{D}} - m)\psi - \frac{1}{4}G^{a\mu\nu}G^a_{\mu\nu}. \tag{1.34}$$

The non-abelian nature of the $SU(3)$ group in QCD results in more complicated terms in the Lagrangian compared to QED. Gluons are *self-coupling* whereas photons do not directly couple to each other.

Electroweak unification

In the first half of the 20th Century, the weak interaction and electromagnetism appeared to be very different phenomena. The weak interaction was first seen in radioactive β decay where a neutron spontaneously decays into a proton, electron and a neutrino ($p \rightarrow ne^-\overline{\nu}_e$) - although it took some time to recognise the presence of the neutrino from the missing energy and momentum.

In a series of developments by Glashow, Salam and Weinberg, the two interactions were unified under a single gauge group - $SU(2)_L \times U(1)_Y$ - where the

subscript L indicates it applies only to left-handed particles and Y is the weak hypercharge [55–57]. The corresponding vector gauge bosons are three W bosons (W^1, W^2, W^3) of weak isospin and the B boson of weak hypercharge. Field strength tensors can be constructed analogously to eqn. (1.27) :

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\epsilon^{ijk} W^j_\mu W^k_\nu, \tag{1.35}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \tag{1.36}$$

The covariant derivative is given by:

$$\mathcal{D}_\mu = \partial_\mu + i\frac{g}{2}\sigma^a W^a_\mu + i\frac{g'}{2}B_\mu. \tag{1.37}$$

The Lagrangian for the fermion sector can now be written:

$$\mathcal{L}_{fermion} = i\overline{\psi}_Q \not{\mathcal{D}} \psi_Q + i\overline{\psi}_{u_R} \not{\mathcal{D}} \psi_{u_R} + i\overline{\psi}_{d_R} \not{\mathcal{D}} \psi_{d_R} + i\overline{\psi}_L \not{\mathcal{D}} \psi_L + i\overline{\psi}_{l_R} \not{\mathcal{D}} \psi_{l_R}. \tag{1.38}$$

Each term corresponds to a row in Table 1.3. First the left-handed quark doublets Q, then the right-handed singlets u_R and d_R; the left-handed lepton doublet L and the right-handed lepton singlet l_R. A right-handed neutrino has not been included as it possesses no electromagnetic, weak or colour charge and participates in no Standard Model interaction. There are current searches for these *sterile neutrinos* [58] but as they are not particularly relevant to this work, they shall not be discussed further.

	Gen. 1	Gen. 2	Gen. 3	I_3	Y
Quarks	$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	$\begin{pmatrix} t \\ b \end{pmatrix}_L$	$1/2$ $-1/2$	$1/3$ $1/3$
	u_R	c_R	t_R	0	$4/3$
	d_R	s_R	b_R	0	$-2/3$
Leptons	$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L$	$1/2$ $-1/2$	-1 -1
	e_R	μ_R	τ_R	0	-2

Table 1.3: Left-handed fermions can be organised into weak-isospin doublets whereas right-handed fermions are singlets.

After EWSB, discussed in Section 1.1.4, the W^a and B fields mix, resulting in the observed $W^\pm$ and Z^0 bosons:

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W^3_\mu \end{pmatrix}, \tag{1.39}$$

$$W^\pm_\mu = \frac{i}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu). \tag{1.40}$$

One can now relate the masses of the $W^\pm$ and Z^0 bosons and also the electric charge in terms of the $SU(2)_L \times U(1)_Y$ couplings and the Weinberg angle θ_W:

$$M_Z = \frac{M_W}{\cos\theta_W}, \tag{1.41}$$

$$e = \frac{g\ g'}{\sqrt{g^2 + g'^2}} = g' \cos\theta_W. \tag{1.42}$$

1.1.4 The Higgs Mechanism

The GSW model unified the electromagnetic and weak interactions into a single $SU(2)_L \times U(1)_Y$ gauge group and derived relations between the gauge fields and the physically observed gauge bosons. There is a hidden problem, however. The physical W and Z bosons are found to have masses of 82 GeV and 91 GeV respectively [59,60] and one encounters a problem when naïvely attempting to include a mass term in the Lagrangian of the form:

$$\mathcal{L}_{\text{naïve}} = m^2 Z^\mu Z_\mu. \tag{1.43}$$

Gauge invariance dictates that the equations of motion derived from the Lagrangian are unchanged under transformations of the type:

$$Z_\mu \to Z_\mu + \partial_\mu \phi, \tag{1.44}$$

where ϕ is an arbitrary scalar function. However, when this transformation is applied to the naïve mass term this gauge invariance is not maintained:

$$\mathcal{L}_{\text{naïve}} \to m^2 Z^\mu Z_\mu + 2m^2 Z^\mu \partial_\mu \phi. \tag{1.45}$$

A similar problem occurs when one writes mass terms for the fermions:

$$m\overline{\psi}\psi = \frac{1}{4} m\overline{\psi}(1-\gamma^5)(1-\gamma^5)\psi + \frac{1}{4} m\overline{\psi}(1+\gamma^5)(1+\gamma^5)\psi \tag{1.46}$$

$$= m\overline{\psi_R}\psi_L + m\overline{\psi_L}\psi_R. \tag{1.47}$$

Right and left-handed fermions have different transformation properties under the $SU(2)_L \times U(1)_Y$ gauge group and such a mass term would break the gauge symmetry.

The solution comes in the form of the Higgs mechanism which posits an $SU(2)$ doublet of complex scalar fields ϕ:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \tag{1.48}$$

which, using the definition of the covariant derivative from eqn. (1.37), contributes the following terms to the Lagrangian:

$$\mathcal{L}_{Higgs} = (\mathcal{D}_\mu \phi)^\dagger (\mathcal{D}^\mu \phi) - \mu^2 |\phi|^2 + \lambda |\phi|^4. \tag{1.49}$$

This is the most general Lagrangian contribution that can be constructed. Lorentz invariance requires even powers of the scalar field, and any contributions that are of order 6 or larger in ϕ are non-renormalisable.

For cases where $\mu^2 > 0$, there is a non-zero Vacuum expectation value (vev), as shown in Figure 1.5 where the minima occurs when the field has a value:

$$\phi^\dagger \phi = \frac{v^2}{2}, \qquad v = \sqrt{\frac{\mu^2}{\lambda}}. \tag{1.50}$$

It is convenient to expand the field around the vev in the *unitary gauge* where

$$\phi \quad = \quad \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{1.51}$$

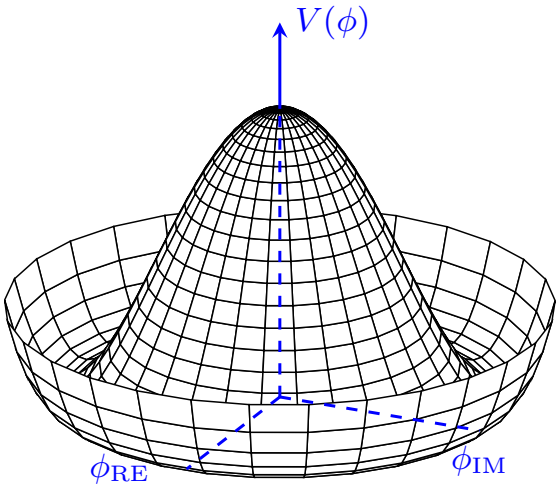


Figure 1.5: The characteristic 'Mexican-hat' shape of the Higgs potential if $\mu^2 > 0$ in eqn. (1.49). The Vacuum expectation value (vev) occurs at the minima of this potential and is non-zero.

The gauge boson masses come from the square of the covariant derivative of the Higgs field, $(D^\mu\phi)^\dagger(D_\mu\phi)$, in the Lagrangian. Expanding out only the relevant terms, one obtains:

$$\begin{aligned}
\Delta\mathcal{L} &= \frac{1}{2}\begin{pmatrix}0 & v\end{pmatrix}\left(\frac{g}{2}W^a_\mu\sigma^a + \frac{g'}{2}B_\mu\right)\left(\frac{g}{2}W^{b\mu}\sigma^b + \frac{g'}{2}B^\mu\right)\begin{pmatrix}0\\ v\end{pmatrix} \\
&= \frac{v^2}{8}\left[\ g^2(W^1_\mu)^2 + g^2(W^2_\mu)^2 + (-gW^3_\mu + g'B_\mu)^2\ \right] \\
&= (\frac{1}{2}vg)^2W^{+\mu}W^-_\mu + \frac{v^2}{8}\begin{pmatrix}B_\mu & W^3_\mu\end{pmatrix}\begin{pmatrix}g'^{\,2} & -gg'\\ -gg' & g^2\end{pmatrix}\begin{pmatrix}B_\mu\\ W^3_\mu\end{pmatrix}. \qquad (1.52)
\end{aligned}$$

The first term is equivalent to the mass term $m_W^2W^+W^-$ with a mass

$$m_W = \frac{vg}{2}. \qquad (1.53)$$

After some trivial but lengthy algebra, the mixing transformation in (1.39) diagonalises the matrix in the second term of (1.52) to

$$\begin{aligned}
&\frac{v^2}{8}\begin{pmatrix}B_\mu & W^3_\mu\end{pmatrix}\begin{pmatrix}g'^{\,2} & -gg'\\ -gg' & g^2\end{pmatrix}\begin{pmatrix}B_\mu\\ W^3_\mu\end{pmatrix} \\
= \frac{v^2}{8(g^2+g'^{\,2})}&\begin{pmatrix}A_\mu & Z_\mu\end{pmatrix}\begin{pmatrix}g & g'\\ -g' & g\end{pmatrix}\begin{pmatrix}g'^{\,2} & -gg'\\ -gg' & g^2\end{pmatrix}\begin{pmatrix}g & -g'\\ g' & g\end{pmatrix}\begin{pmatrix}A_\mu\\ Z_\mu\end{pmatrix} \\
&= \frac{v^2}{8}\begin{pmatrix}A_\mu & Z_\mu\end{pmatrix}\begin{pmatrix}0 & 0\\ 0 & g^2+g'^{\,2}\end{pmatrix}\begin{pmatrix}A_\mu\\ Z_\mu\end{pmatrix}. \qquad (1.54)
\end{aligned}$$

To be clear, this is not just mathematical trickery. The extra terms in the Lagrangian can be interpreted only as mass terms if the matrix is diagonal and by diagonalising the matrix, one recovers the physical photon and Z^0 boson field. The masses can be read off from the diagonal elements of the matrix, which gives a massless photon and a Z^0 boson with mass:

$$m_Z = \frac{v}{2}\sqrt{g^2+g'^2}. \qquad (1.55)$$

1.1.5 Masses of the fermions

The previous section details how the Higgs mechanism can generate masses for the gauge bosons, but the masses of the fermions can now be re-interpreted as their couplings to the Higgs field. For instance, terms coupling the quarks to the Higgs field can be written as:

$$\mathcal{L}_{quark} = -Y_d\overline{d_L}\phi d_R - Y_u\overline{u_L}\phi u_R \tag{1.56}$$

$$= -\frac{v}{2}\left[Y_d\overline{d_L}d_R + Y_u\overline{u_L}u_R\right] - \frac{h}{2}\left[Y_d\overline{d_L}d_R + Y_u\overline{u_L}u_R\right], \tag{1.57}$$

where the arbitrary constants $Y_{u,d}$ are the coupling strength to the Higgs field and cannot be predicted in the Standard Model. Terms involving the vev are identical to mass terms in the Lagrangian with a mass given by $m_f = \frac{Y_f v}{2}$. One small subtlety is that the mass matrices for up-type and down-type quarks cannot be simultaneously diagonalised. This in turn means that mass eigenstates - those that propagate freely - are not the same as weak-eigenstates - those involved in weak interactions.

The transformation between these eigenstates is encoded in the Cabibbo-–Kobayashi-–Maskawa (CKM) and Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrices [61, 62]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_{Weak} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{Mass} \tag{1.58}$$

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix}_{Weak} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix}_{PMNS} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}_{Mass}. \tag{1.59}$$

1.2 Problems with the Standard Model

1.2.1 Gravity

Perhaps the most glaring omission in the Standard Model is that it does not include gravity. The most successful current description of gravity is given by Einstein's theory of general relativity [63]. An important prediction of the theory is the existence of gravitational waves which were directly detected in September 2015 by the LIGO and Virgo collaborations [64]. General relativity interprets gravity as the curvature of spacetime which is dictated by the energy and momentum of matter and radiation according to Einstein's field equations:

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}. \tag{1.60}$$

The left-side describes the curvature of spacetime via the Ricci scalar and tensor (R and $R_{\mu\nu}$) and the right hand side contains the mass-energy-momentum tensor $T_{\mu\nu}$. The constant of proportionality contains the gravitational constant and the speed of light in vacuum and its small magnitude, $\mathcal{O}(10^{-43}\ \mathrm{s^2 m^{-1} kg^{-1}})$, explains why the curvature even around massive objects like the Earth can still be quite small.

The solutions of eqn. (1.60) are spacetime metrics like the Minkowski and Schwarzschild metrics:

$$Minkowski \quad g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{1.61}$$

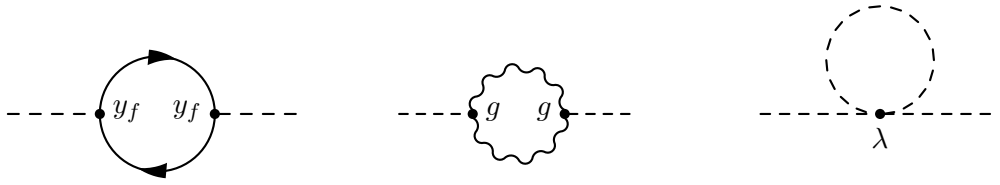
$$Schwarzschild \quad g_{\mu\nu} = \begin{pmatrix} (1 - \frac{\mu}{r}) & 0 & 0 & 0 \\ 0 & -\frac{1}{(1-\frac{\mu}{r})} & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2 \sin^2\theta \end{pmatrix} \quad \text{where} \quad \mu = \frac{2GM}{c^2}. \tag{1.62}$$

The Standard Model uses the Minkowski metric for Euclidean (flat) spacetime and whilst in most cases this is a reasonable approximation, it is clear that it is not valid in regions of large curvature, such as in the centre of black holes.

Since electromagnetism can be described using a classical field theory and then a corresponding Quantum Field Theory was found, one might hope that the classical field theory of gravity could be developed in the same way. Attempts have been made to quantise the tensor field in eqn. (1.60), which gives rise to a spin-2 mediator, the graviton [65]. Unfortunately such theories are non-renormalisable (see Section 1.1.3) and therefore not useful descriptions of physical phenomena. It is widely believed that a new theoretical framework, such as string theory [66] or quantum-loop gravity [67] is needed to describe gravity in a consistent way with the other fundamental interactions in the Universe.

1.2.2 Naturalness

An accurate calculation of the physical mass of the Higgs boson requires the inclusion of all possible radiative corrections. Consider the contribution of diagrams such as Figure 1.6



(a) (b) (c)

Figure 1.6: Example Feynman diagrams for contributions to the radiative corrections to the Higgs mass in the Standard Model from a fermion loop (a), a boson loop, (b) and from the Higgs self interaction (c).

These introduce corrections to the mass of the form

$$\delta m^2 = \frac{1}{16\pi^2}\left[-6y_f^2 + g^2 + 6\lambda\right]\Lambda^2, \tag{1.63}$$

where Λ is the upper-limit of the integration over particle momenta known as the UV cut-off. As Λ becomes larger this correction diverges. This problem can be dealt with using renormalisation, where the physical mass of the Higgs boson is not the bare mass which appears in the Lagrangian. However, if one believes that the Standard Model is valid up to the energy scale of the Planck mass ($M_p \approx 2 \times 10^{18}$ GeV), this requires the cancellation of very large terms which result in a mass of only $\mathcal{O}(100$ GeV). This is sometimes also known as the *fine-tuning* problem [51, p. 788].

1.2.3 Dark Matter

For many years it has been understood that the visible baryonic matter in the universe comprises only a small fraction of the total mass-energy in the Universe. The current best explanation for this discrepancy is that there exists a new form of matter which interacts gravitationally but does not interact with light - hence the name 'Dark Matter'.

In 1937 Fritz Zwicky studied the Coma galaxy cluster, using the Virial theorem to estimate the total mass of the cluster from the velocities of the constituent galaxies [68]. This was compared with the total luminosity of the cluster to find the average mass-to-light ratio 800 times larger than the mass-to-light ratio of the Sun.

This seems inconsistent with a galaxy mostly composed of stars similar to the Sun suggesting much of the galaxy's mass is not luminous.

In 1980, a similar study was performed for individual galaxies, looking at the 'rotation-curves' (see Figure 1.7a) which give the distribution of stellar velocities as a function of their distance from the galactic centre. If the mass of the galaxy is concentrated at the centre, as the distribution of light seems to suggest, then one would expect the average stellar velocity to decrease further from the centre. The observed distributions are actually flat, which implies that a large amount of the mass in the galaxy is at a larger radial distance [69].

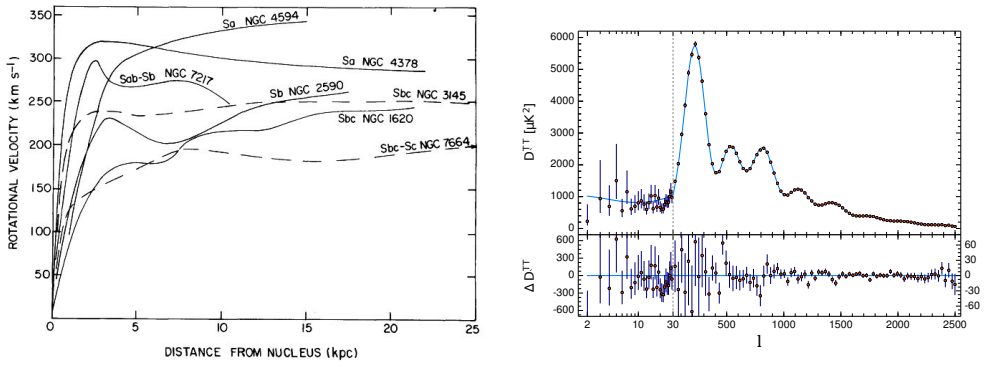
Another source of evidence for the existence of Dark Matter comes from the fluctuations in the Cosmic Microwave Background Radiation (CMBR), which has been measured with incredible precision by the WMAP and Planck space telescopes. The fluctuations can be decomposed into their harmonic components, a series of Legendre polynomials as shown in Figure 1.7b. This distribution is highly sensitive to the matter content of the Universe and a fit to this data can provide an estimate of the missing mass content. The most recent Planck result calculates that dark matter comprises 84% of the total matter content of the Universe [70]. The distribution of Dark Matter can be inferred from gravitational lensing, where the path of light is curved by the dark matter between the source and the observer.

There are no Standard Model candidates for Dark Matter because:

- Dark matter cannot couple to the electromagnetic interaction because otherwise it would be visible. This eliminates all particles with charge.
- As Dark matter remains in the Universe, it must be stable - at least over cosmological timescales.
- The distribution of Dark Matter suggests it is 'cold' - it is at low-enough velocities that it coalesces. This eliminates the neutrinos which have always been observed to be highly-relativistic.

1.2.4 Matter-antimatter Asymmetry

In our everyday experience, we see only matter and no antimatter, whilst most physical processes in the Standard Model seem to have no preference between the two. If large amounts of antimatter were present, one would expect to observe the ongoing annihilation at the boundary with ordinary matter. Since these annihilations have not been observed, it is natural to infer that there is an overall asymmetry in the visible Universe.



(a) (b)

Figure 1.7: (a) Rotational velocities for stars associated to seven galaxies, the plateau at larger distances implies the existence of dark matter in a halo around each galaxy [71]. (b) Evidence for dark matter in the power spectrum of the Cosmic Microwave Background measured by the Planck Collaboration [70].

In 1967 Andrei Sakharov discussed the conditions necessary for baryon Asymmetry that have come to be known as the *Sakharov conditions* [72]:

i Baryon Number violation

ii C and CP violation

iii Interactions occur out of thermal equilibrium.

The requirement for C and CP violation is necessary as for any process that produces a baryon (e.g $X \to Y + B$) there are conjugate processes ($\overline{X} \to \overline{Y} + \overline{B}$) that produce an anti-baryon. These processes must occur at different rates for *baryogenesis* to take place. Similarly, the third requirement is necessary to prevent the reverse process $Y + B \to X$ from cancelling out the baryogenesis. In cosmological terms, this requires that the rate of a reaction that generates baryon-asymmetry must be less than the rate of expansion of the Universe.

1.3 Beyond the Standard Model (BSM)

1.3.1 Supersymmetry

Supersymmetry (SUSY) is a proposed symmetry of nature where every Standard Model fermion has a supersymmetric partner boson, and likewise each SM boson has a SUSY fermion. This symmetry has many interesting consequences, some of which could provide solutions to some of the problems with the Standard Model discussed

in Section 1.2. Firstly, bosonic fields give positive contributions to the vacuum energy, which arises due to fluctuations in the quantum fields in the vacumm, whilst fermionic fields give negative contributions. Therefore when one includes the effect of supersymmetric particles in renormalisation theory, one finds that many divergences cancel exactly - to all orders of perturbation theory. In such a scenario, the Higgs mass would not need huge corrections and the naturalness problem (Section 1.2.2) can be mitigated [73].

In supersymmetric theories, Baryon and Lepton numbers are not conserved and this would allow proton decay to occur at a rate higher than the experimental limits. Therefore a new discrete symmetry is proposed called R-parity which is defined as [74]

$$P_R = (-1)^{3B+L+2s}, \tag{1.64}$$

where B and L are the Baryon and Lepton numbers respectively, and s is spin. Standard Model particles have R-parity $(+1)$ and supersymmetric particles have R-parity (-1). If this quantity were conserved, it would prevent proton decay but it would also imply that supersymmetric particles would all eventually decay into the Lightest Stable Particle (LSP), which cannot decay into Standard Model particles. This LSP would exist as a cosmic relic, and if were of the right mass, it could be a very promising candidate for Dark Matter [75].

Since there is no experimental evidence for a charged boson with the same mass as the electron, Supersymmetry cannot be an exact symmetry of the Universe, however it could be a spontaneously broken symmetry, just as in the case of EWSB. In this case, the masses of the supersymmetric particles could be larger by $\mathcal{O}(100\text{ GeV})$ whilst still retaining all of the useful properties discussed previously.

SUSY introduces as many as 128 free parameters such as the masses and couplings of the new particles. It is very difficult to explore such a large parameter space experimentally. A number of models, such as the MSSM are a subset of SUSY which make some simplifying assumptions about the large parameter space. This provides a benchmark that experimentalists can use in their searches [76].

The MSSM predicts a partner to the Standard Model's Higgs doublet and is therefore an example of a 2HDM. As discussed in Section 1.1.5, the up and down-type fermion mass matrices cannot be simultaneously diagonalised. In the MSSM this implies that the two Higgs doublets provide mass to the up and down-type fermions separately. Therefore, one can write the doublets as

$$H_u = \begin{pmatrix} \phi_u^+ \\ \frac{1}{\sqrt{2}}\left(v_u + \phi_u^0 + i\chi_u^0\right) \end{pmatrix}, \quad H_d = \begin{pmatrix} \frac{1}{\sqrt{2}}\left(v_d + \phi_d^0 + i\chi_d^0\right) \\ \phi_d^- \end{pmatrix}, \tag{1.65}$$

where $v_u + v_d = v_W = (246\text{ GeV})^2$ is the SM Higgs vev and the ratio of the two vevs is $\tan(\beta) = \frac{v_u}{v_d}$. After EWSB, where the gauge bosons acquire mass, there are 5 remaining bosons: 2 CP-even neutral bosons h^0, H^0; 1 CP-odd neutral boson A^0 and two charged bosons $H^\pm$

$$\begin{pmatrix} H^0 \\ h^0 \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \phi_u^0 \\ \phi_d^0 \end{pmatrix}, \quad \begin{pmatrix} G^0 \\ A^0 \end{pmatrix} = \begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} \chi_u^0 \\ \chi_d^0 \end{pmatrix}, \tag{1.66}$$

$$\begin{pmatrix} G^\pm \\ H^\pm \end{pmatrix} = \begin{pmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} \phi_u^\pm \\ \phi_d^\pm \end{pmatrix}. \tag{1.67}$$

The mixing angle α is the angle required to diagonalise the mass matrix for h^0 and H^0. There are only two free parameters in the Higgs sector of the MSSM, conventionally taken to be $\tan\beta$ and M_A. Searches for the MSSM Higgs bosons must therefore search the $\tan\beta$ vs. M_A plane. The couplings of the heavy Higgs bosons to down-type fermions is greatly enhanced for large $\tan\beta$ values. In this case the cross-section of the b-associated production mechanism, shown in Figure 1.8a can become as large as the gluon-fusion production mechanism (Figure 1.8b). The associated b-quarks can be identified in a process called *b-Tagging*, described in detail in Section 3.4.

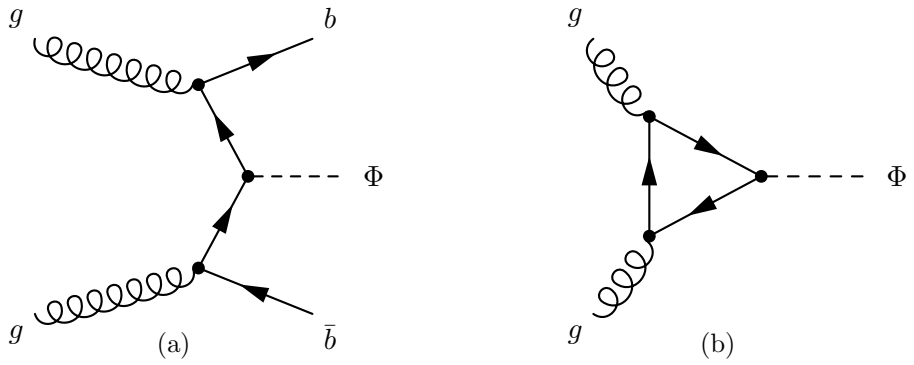
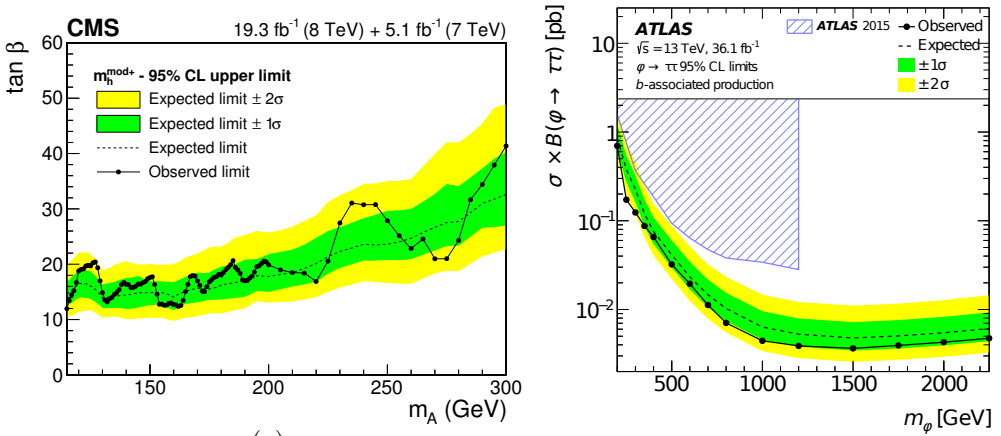


Figure 1.8: Feynman diagrams for a b-quark associated production mode and b gluon-gluon fusion production mode. The scalar Φ represents any of the neutral scalars in the MSSM h^0, H^0, A^0.

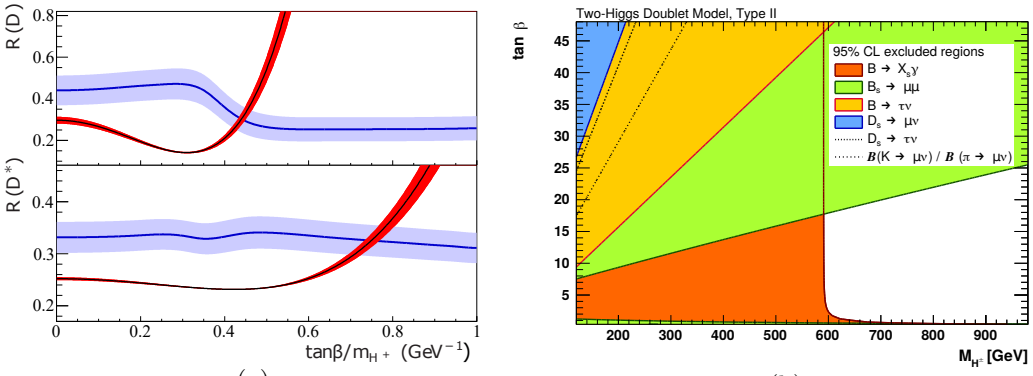
Previous searches for these heavy Higgs bosons have been performed, in the dimuon channel in LHC Run 1 [77] and in the ditau channel in Run 2 [78]. The CMS dimuon analysis placed upper limits on the $\tan\beta$ as a function of m_A up to a maximum mass of 300 GeV. The ATLAS ditau search placed limits on the cross-section times branching ratio for the b-associated production of a heavy Higgs decaying to a pair of tau particles up to a mass of 2.25 TeV. These results are shown in Figure 1.9.



(a) (b)

Figure 1.9: (a) Limits on $\tan\beta$ as a function of m_A from a CMS Run 1 search in the dimuon decay channel [77]. (b) Cross-section times branching ratio limits for b-associated production of a heavy Higgs with ATLAS 13 TeV data in the ditau decay channel [78]. Areas above the solid, observed, limits corresponding to large $\tan\beta$ and large cross-sections, have been excluded.

There are also important constraints on the MSSM from flavour physics experiments. If there were extra Higgs doublets that coupled to muons, one would expect this to affect the rate of the rare decay, $B_s \to \mu^+\mu^-$. However in a recent measurement by the LHCb experiment, no significant deviations from the Standard Model were measured [79]. Similarly, one can search for the charged Higgs bosons arising from the MSSM. There are strong constraints from analyses of the rate of $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ decays which completely exclude a class of 2HDMs, known as 'Type II', where up- and down- type quarks couple to different doublets [16, 18, 21]. Figure 1.10a shows the results of measurements by the BaBar collaboration of two ratios, $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$, defined as $\frac{BR(\overline{B}\to D^{(*)}\tau\overline{\nu}_\tau)}{BR(\overline{B}\to D^{(*)}l\overline{\nu}_l)}$ where l is either a muon or electron. One can see that a consistent value of $\tan\beta/\mathrm{m}_{H^+}$ cannot be found for both measurements, and the analysis found that this type of model can be excluded to the 99.8% confidence level [18]. Furthermore, the branching ratio of $B \to \tau\nu$ is also sensitive to the presence of charged Higgs bosons. In a global fit of electroweak parameters [80], using an average of $BR(B \to \tau\nu)$ measurements from the Heavy Flavour Averaging Group [81], strong constraints in the $\tan\beta$-m_{H^+} plane are found, as shown in Figure 1.10b.



(a) (b)

Figure 1.10: (a) Comparison of the results of the $\overline{B} \to D^{(*)}\tau^-\overline{\nu}_\tau$ analysis from the BaBar collaboration (blue) with predictions including a charged Higgs boson of a type II 2HDM model (red) shown as a function of $\tan\beta/\mathrm{m}_{H^+}$. The Standard Model corresponds to $\tan\beta/\mathrm{m}_{H^+} = 0$. The ratios $\mathcal{R}(D^{(*)})$ are defined as $\frac{BR(\overline{B}\to D^{(*)}\tau\overline{\nu}_\tau)}{BR(\overline{B}\to D^{(*)}l\overline{\nu}_l)}$ where l is either a muon or electron [18]. (b) Exclusion limits in the $\tan\beta$-m_{H^+} plane for a type II 2HDM model using constraints from flavour observables [80].

1.3.2 Flavourful Higgs

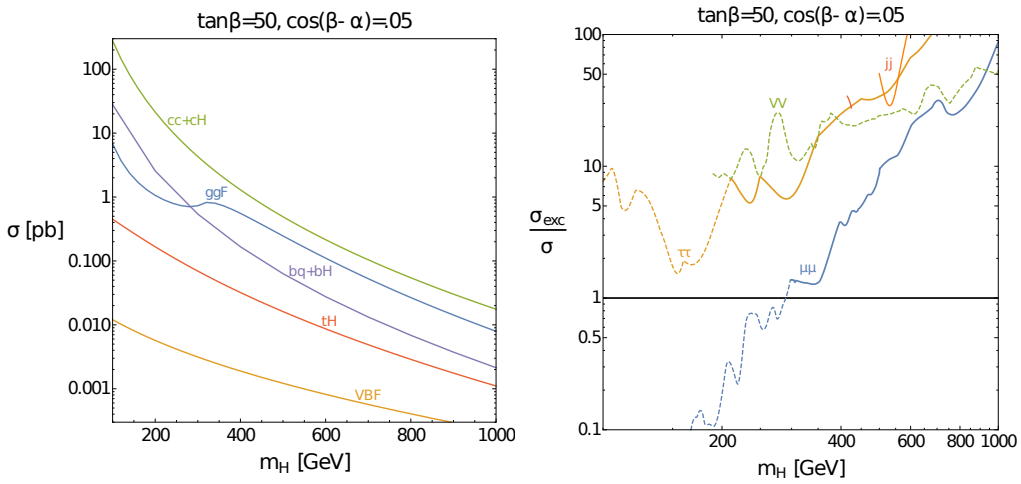
As of yet, the only observed decays of the 125 GeV Higgs boson to fermions are in the $H \to \tau^+\tau^-$ and $H \to b\bar{b}$ channels, both of which are 3^{rd} generation fermions. Some extensions to the Standard Model allow extra Higgs doublets to couple differently between the generations. One example of a *Flavourful Higgs* model posits two Higgs doublets, each with weak hypercharge +1/2 [82]

$$\Phi = \begin{pmatrix} \phi^+ \\ \frac{1}{\sqrt{2}}(v + \phi + ia) \end{pmatrix}, \Phi' = \begin{pmatrix} \phi'^+ \\ \frac{1}{\sqrt{2}}(v' + \phi' + ia') \end{pmatrix}, \tag{1.68}$$

where $v + v' = v_W = (246\ \mathrm{GeV})^2$ is the SM Higgs vev and the ratio of the two vevs is $\tan(\beta) = \frac{v}{v'}$. The mass matrices of the charged SM fermions receive contributions from both Higgs doublets:

$$m^{(\prime)}_{xx'} = \frac{v^{(\prime)}}{\sqrt{2}} \langle x_L | \lambda^{(\prime)x} | x'_R \rangle, \quad \text{for } x \in \{u, c, t, d, s, b, e, \mu, \tau\}. \tag{1.69}$$

These mass matrices are the equivalent of the SM matrices in eqn. (1.56), although there are now two vevs, v and v' and two sets of coupling constants, written in the form of so-called *Yukawa textures*, λ and λ'. These Yukawa textures could be flavour-dependent however, and the authors of this model consider a lepton Yukawa texture for which the doublet, Φ only couples to the third generation [82]:



(a) (b)

Figure 1.11: (a) Production cross-sections for neutral scalar H as function of the mass m_H for different production mechanisms. (b) Exclusion limits for the heavy scalar boson, normalised to the predicted cross-section, from previous searches [82].

$$\lambda^l \sim \frac{\sqrt{2}}{v}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & m_\tau \end{pmatrix}, \qquad \lambda'^l \sim \frac{\sqrt{2}}{v'}\begin{pmatrix} m_e & m_e & m_e \\ m_e & m_\mu & m_\mu \\ m_e & m_\mu & m_\mu \end{pmatrix}. \tag{1.70}$$

In this Flavourful Higgs model, just as for the MSSM, the dominant production mechanisms are dependent on $\tan\beta$, and in certain cases, the c-associated and b-associated mechanisms have comparable cross-sections to the gluon-fusion mechanism as one can see in Figure 1.11a. The authors of this model have used previous searches to exclude regions of the parameter space and found that the most stringent limits came from the dimuon decay channel. However, Figure 1.11b shows that they were only able to exclude their model for Higgs masses $m_H \lessapprox 360$ GeV. The dimuon channel remains the most sensitive up to masses of 1 TeV.

1.3.3 Z-prime

The Standard Model already includes a $U(1)$ gauge symmetry which results in the photon. It is therefore reasonable to posit that additional $U(1)$ symmetries may exist which, after symmetry breaking, would result in an additional massive gauge boson. These are often called Z′ in analogy with the Standard Model gauge bosons.

Recent searches for such a particle, decaying into two muons or two electrons have been performed at the LHC by the ATLAS and CMS experiments [83, 84]. No evidence for a new particle has yet been discovered, and the ATLAS search excluded,

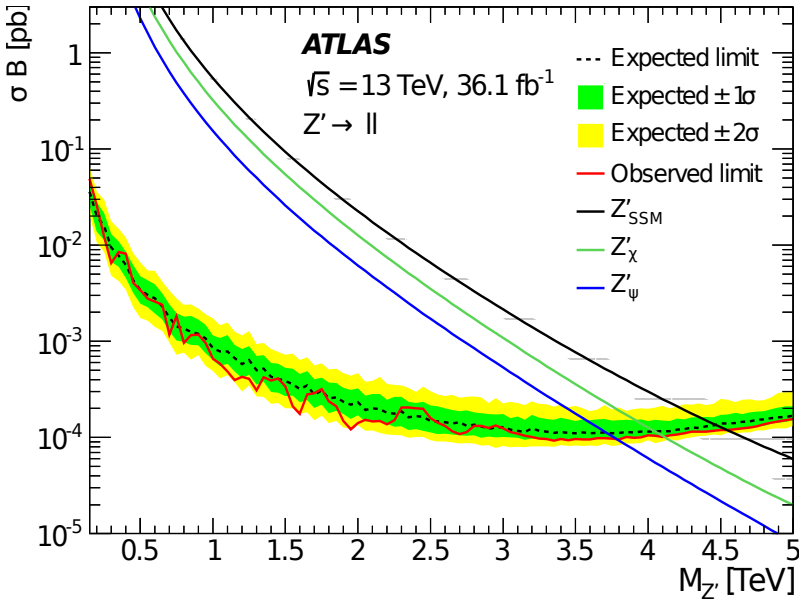


Figure 1.12: Recent results of an ATLAS search for a Z′ boson decaying into electron-positron or muon-antimuon pairs. The three models they consider are excluded for a Z′ mass of below 3.8 TeV [83].

to 95% confidence level (see Section 6), models which include a Z′ with a mass below 3.8 TeV as shown in Figure 1.12.

1.3.4 Grand Unification

The unification of the electromagnetic and weak forces was described in Section 1.1.3. It is only natural to consider if the strong force can also be unified in this way resulting in a Grand Unified Theory (GUT). Figure 1.13 shows the evolution of the gauge couplings to high energies and at a scale of $\mathcal{O}(10^{15}$ GeV) they take similar values. In fact, if one includes the effects of Supersymmetry this crossing becomes more exact. This suggests that there may be some larger gauge symmetry with a single unified coupling constant, but at energies below the GUT scale (Λ_{GUT}), the symmetry is broken resulting in the $SU(3) \times SU(2) \times U(1)$ structure of the Standard Model. An early example was the $SU(5)$ Georgi-Glashow model [85] which has since been well excluded by experimental limits on the proton lifetime [86].

1.3.5 Extra dimensions

In our everyday experience, we observe the 3 dimensions of space and 1 dimension of time that Einstein described in his theory of special relativity as a single entity - spacetime. It is possible, however, that extra dimensions exist but are contracted

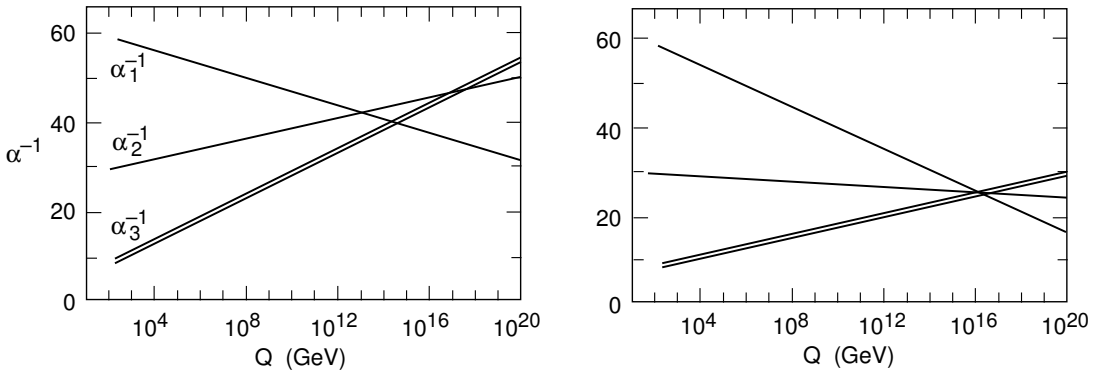


Figure 1.13: Evolution of $SU(3) \times SU(2) \times U(1)$ gauge couplings to high energy scales using the Standard Model renormalisation group equations (a) and including Supersymmetry (b) [87] where α_1, α_2 and α_3 correspond to the $U(1)$, $SU(2)$ and $SU(3)$ gauge couplings respectively.

to very small distance scales. This could explain the relative weakness of gravity; if gravity acted throughout each of the dimensions and the other fundamental interactions only operated in $(3+1)$ dimensions, they would appear to have different strengths [88].

Experimental searches for extra dimensions often focus on searching for new high-mass particles. At first glance it may appear odd that additional dimensions would result in new particles as no new symmetry has been introduced. In fact, they would not be new 'particles' *per se*, rather they would be heavy 'states' of the particles in the Standard Model. This outcome was first noted in Kaluza-Klein theory, which attempted to unify electromagnetism and gravity in a five-dimensional spacetime [89, 90]. The relativistic relation between energy and momentum is

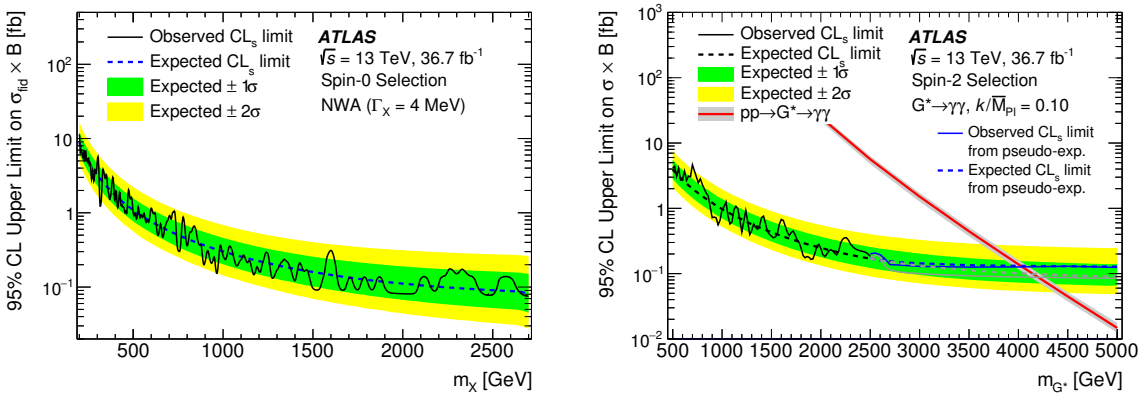
$$E^2 = |\vec{p}|^2 + m^2 = p_x^2 + p_y^2 + p_z^2 + m^2. \tag{1.71}$$

If there were another dimension, the momentum in that direction would appear as a larger mass term:

$$E^2 = p_x^2 + p_y^2 + p_z^2 + \left[p_{extra}^2 + m^2\right] \tag{1.72}$$

$$= p_x^2 + p_y^2 + p_z^2 + \left[M^2\right]. \tag{1.73}$$

As the size of this extra dimension, R_5, would be finite and small, the particle's wavelength can only take certain discrete values and so, via the de Broglie equation, the momentum in this direction would also be quantised in units of $\frac{\hbar}{R_5}$. Therefore, the existence of extra dimensions would result in high-mass partner states - known



(a) (b)

Figure 1.14: Cross-section times Branching Ratio limits on diphotons produced by a spin-0 resonance (a) and the lightest Kaluza-Klein graviton, which would be spin-2 (b) [91].

as Kaluza-Klein resonances - of the Standard Model particles. If discovered, the mass of these states would allow the calculation of the size of this extra dimension.

A search for Kaluza-Klein resonances decaying into two photons at the ATLAS detector was performed for Run 2 (13 TeV) data [91]. The cross-section times branching ratio limits are shown in Figure 1.14; no significant excesses of data were observed when searching for spin-0 or spin-2 resonances. They exclude a particular extra-dimensional model, the Randall-Sundrum (RS1) model [92] with a Kaluza-Klein graviton mass below ($m_{G^*} = 4.1$ TeV).

1.4 Motivation for search for dimuon resonance + b-jet

As discussed in the previous sections, there are several outstanding questions regarding the Standard Model and a multitude of models which extend it. A search for 'New Physics' often entails searching for the existence of new particles and this requires understanding the signatures that these new particles will leave in detectors. In the case of the MSSM, Flavourful Higgs and similar models which predict that there may be more than one Higgs doublet involved in the EWSB process, the extra bosons introduced would decay into, among other things, a muon-antimuon pair. The telltale signature would be the resonance peak in the dimuon mass spectrum at the mass of this new particle.

In many models the expected Branching Ratio (BR) - the fraction of decays into a particular channel - for muons is very small. Consider the case of the MSSM,

where the couplings of fermions to the heavy Higgs bosons are proportional to their mass. In this case the decay to the tau lepton is approximately $\frac{BR(\Phi\to\tau\tau)}{BR(\Phi\to\mu\mu)} \approx \mathcal{O}(\left(\frac{m_\tau}{m_\mu}\right)^2) \approx 300$ times more likely than for the muon. The dimuon channel has its advantages, however, as muons are more easily identified and measured in general purpose detectors like ATLAS.
The decay of the 125 GeV boson to muons has yet to be discovered [93], although it is expected that enough data to reach sensitivity will be collected by the end of Run 3 of the LHC in 2023. This is an important, and in the author's opinion exciting, physics goal for the LHC. Since only the 3^{rd} generation fermion couplings to the 125 GeV boson have been measured, this leaves the possibility that the masses of the other fermions are generated by another Higgs doublet as in the Flavourful Higgs model. In this case, the best channels to search for are the dimuon ($\mu^+\mu^-$) and dicharm ($c\bar{c}$) decays. The dicharm signature suffers from a very large background in hadron colliders, however.

As mentioned in Section 1.3.1, the $\tan\beta$ parameter describes the ratio of the vevs in a 2HDM model, and in the case that $\tan\beta$ is large, the couplings to down-type fermions are greatly enhanced. In such a scenario the b-associated production mechanism (Figure 1.8a) can represent a significant fraction of the heavy Higgs production cross-section. Searching for this production mechanism often adds sensitivity as b-jets can be reasonably well identified and a good signal-to-background ratio can be obtained.

This motivates a search for a resonance in the dimuon spectrum, at masses higher than those explored by the SM $H \to \mu\mu$ analysis [93], with an associated b-jet. As described, there are many models which would give interesting signatures, and so it was decided that the most useful results would be as *model independent* as possible. Rather than interpreting results in a particular model, it is valuable to allow theorists to interpret the outcomes of the search in any model they choose. As such, the objective for this search was to find and measure, or place exclusion limits on, the Cross-section times Branching Ratio ($\sigma_\Phi \times \mathcal{B}(\Phi \to \mu\mu)$) of a model-independent $\Phi \to \mu\mu$ process with and without associated b-jets.

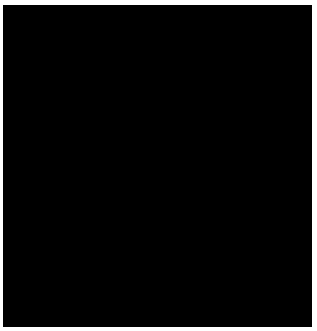
The methodology and results of such an analysis, performed on LHC Run 2 data at the ATLAS experiment, form a large part of the remainder of this thesis.

Chapter 2

The ATLAS Experiment

It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong.

Richard P. Feynman



2.1 The Large Hadron Collider

The LHC is the largest particle collider in the world at 27 km in circumference and is housed in a circular tunnel buried approximately 100 m under the French-Swiss border near Geneva [94–96]. It is designed to provide the highest centre-of-mass energies at 14 TeV and is therefore at the forefront of high-energy particle physics technologies. There are four experiments at different interaction points around the ring: ALICE, LHCb, CMS and the most important - for the purposes of this thesis - ATLAS. This chapter will start with a summary of some key concepts in proton-proton collision physics and then continue with the specifics of the proton beams collided at the LHC. The latter half of this chapter is concerned with the ATLAS

detector, beginning with a summary of important principles in particle detection and then describing the design and performance of the detector itself.

2.1.1 Physics of proton collisions

Hadrons are bound states containing many partons and it is these partons which actually interact during a proton-proton collision. It is therefore not possible to know the exact momentum of an individual parton. A number of assumptions can be made to simplify the problem, however. When considering *hard* scattering processes, where the energy scale is well above ≈ 200 MeV and QCD is perturbative, the partons can be considered as free particles throughout the interaction. The momentum of each parton, p_i, is a fraction of the total proton momentum, p_{Total} and this fraction is called the Bjorken x [97]:

$$x_i = \frac{p_i}{p_{\text{Total}}}. \tag{2.1}$$

It is convenient to separate the proton-proton cross-section into terms which contain the parton content of each proton and the hard-scattering of the partons themselves as

$$\sigma(p + p \to X) = \sum_{ij} \int_0^1 dx_i \int_0^1 dx_j f(x_i) f(x_j) \times \sigma(q_i + q_j \to X). \tag{2.2}$$

This factorisation is illustrated in Figure 2.1. The partonic cross-section ($\sigma(q_i + q_j \to X)$) can be calculated following the prescription in Section 1.1.2. The functions $f(x_i)$ are known as PDFs which give the probability of finding a parton with a given momentum fraction. They cannot be calculated *a priori* using perturbative QCD and are instead found by fitting observables to experimental data (see Section 4.1.1 for more details) .

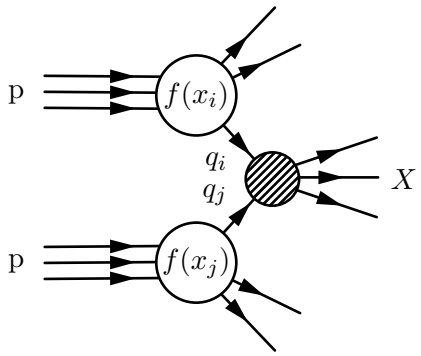


Figure 2.1: The contributions of the hard interaction and the parton distribution function to the overall cross-section can be factorised.

Luminosity

The number of events in a collider experiment can be related to the cross-section of the relevant physics process via the luminosity as

$$N_{event} = \sigma_{event} L = \sigma_{event} \int \mathcal{L} dt. \tag{2.3}$$

The integrated luminosity, L, represents the size of a dataset and is conventionally given in units of *inverse femtobarns* (fb^{-1}). The instantaneous luminosity, $\mathcal{L}$, is a measure of the proton collision rate and is given by [98]

$$\mathcal{L} = \frac{N_1 N_2 f_{rev} N_b}{4\pi\sigma_x\sigma_y} F, \tag{2.4}$$

where N_1 & N_2 are the number of protons per bunch, N_b is the number of bunches, f_{rev} is the revolution frequency, σ_x & σ_y are the widths of the Gaussian beam profile and F is a factor that depends on the crossing angle of the beam. In order to maximise the number of interesting physics events, it is necessary to maximise the instantaneous luminosity. The design luminosity for the LHC is 1×10^{34} $\text{cm}^{-2}\text{s}^{-1}$ although for much of Run 2 approximately double this luminosity was achieved [99].

The measurement of the luminosity requires two components: dedicated detectors known as luminometers measure a *visible* interaction rate, μ_{vis}. Correction factors to calculate the *total* interaction rate from μ_{vis} are determined by performing dedicated Van der Meer (VdM) scans [100]. During a VdM scan, the beams start with some transverse separation and then one beam is 'swept' through the other. Two scans are performed in orthogonal directions to determine the beam profile σ_x and σ_y, and then eqn. (2.4) is used to calculate the total instantaneous luminosity. The necessary correction factor is then simply, $\epsilon = \frac{\mu_{\text{vis}}}{\mu_{\text{total}}}$ and can be used to determine the total instantaneous luminosity during normal physics data-taking runs [100].

2.1.2 Proton beam production

The details of how the LHC beam is produced and manipulated could fill this thesis many times over. Instead, this section will include a brief summary to give the reader context. In simple terms, acceleration is achieved by the use of Radio Frequency (RF) resonant cavities where the electric field is modulated and then dipole magnets are used to direct the beam around in its circuit. Quadrupole and octupole magnets are used to focus the beam towards the interaction points.

Accelerators

In order to reach high enough energies, protons are accelerated through a series of progressively more powerful accelerators before being injected into the LHC as shown in Figure 2.2 [96]. At the start of this 'injection chain', hydrogen atoms are ionised, separating the protons and electrons. A so-called *bunch* containing approximately 10^{11} of these protons is then injected into a linear accelerator called the Alvarez Proton Linac, also known as Linac 2, and then injected into a series of increasingly large synchrotrons. Firstly the Proton Synchrotron Booster (PSB) accelerates the protons from 50 MeV to 1.4 GeV before injecting them into the Proton Synchrotron (PS) which is designed to accelerate protons to an energy of 25 GeV. The proton bunches are then accelerated by the Super Proton Synchrotron (SPS) and finally the LHC which reach energies of 450 GeV and 6.5 TeV respectively. The accelerators used for the LHC are summarised in Table 2.1.

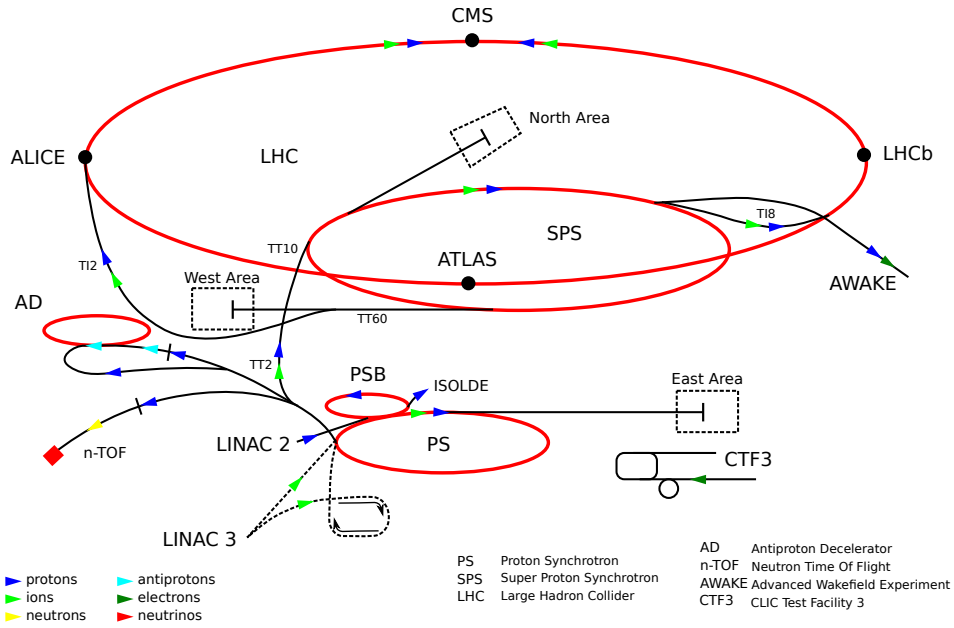


Figure 2.2: The injector chain for the LHC. Protons are injected using Linac 2 into the PSB. After this the protons are accelerated by the PS, the SPS and then the LHC [101].

Bunch structure

As suggested by the name, synchrotrons require that the rate a particle circumnavigates the accelerator is synchronised with the frequency of the voltage oscillations in the RF cavities. Therefore, there are only a given number of positions that a

Table 2.1: Summary of accelerators used for hadron beams at the LHC [96].

Accelerator	**Radius [km]**	**Final Energy [GeV]**
Linac 2	N.A	0.05
PSB	0.025	1.4
PS	0.072	25
SPS	1.1	450
LHC	26.7	6500

particle can take to ensure it is always accelerated by the oscillating field. In the LHC, where the RF cavities operate at 400.8 MHz , there are approximately 36,000 of these locations, known as *buckets*. However, during normal operation, only about 2,800 of these buckets are filled by proton bunches. The reason for this is twofold: firstly, sets of 10 buckets are each assigned a Bunch Crossing Identifier (BCID) which contain only one bunch. BCIDs are separated by 10 buckets to minimise out-of-time pileup and reduce the load on detector readout electronics. Another reason for empty buckets is to leave an *abort gap*. At the end of a run or if a problem occurs, it takes a short but non-negligible time to adjust the magnets so that the beam safely leaves the LHC and is *dumped*. It would be dangerous for proton bunches to be in this section of the beampipe whilst this transition is taking place.

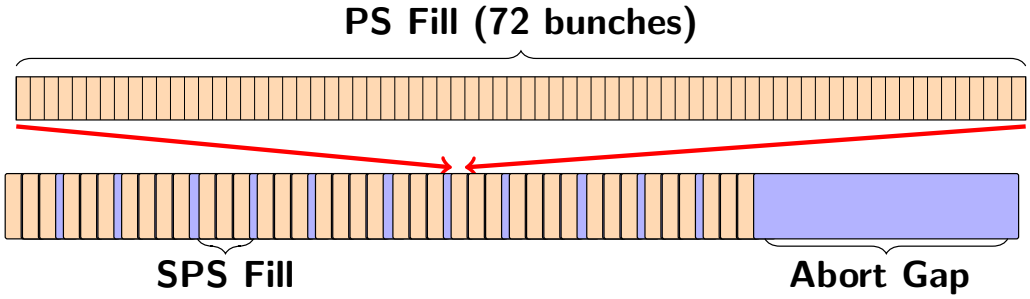


Figure 2.3: Schematic of nominal bunch structure for proton collisions at the LHC.

The exact structure of bunches is often adjusted between different LHC fills to ensure beam stability but a nominal fill contains 2808 bunches. The PS accelerator can hold 72 bunches, which is also known as a *bunch train*. A series of bunch trains are injected into the SPS separated by eight empty BCIDs. A nominal LHC fill comprises 39 bunch trains - separated by small gaps - along with a larger, 1128 BCID abort gap, as shown in Figure 2.3. The choice of LHC filling scheme affects each of the experiments as they need to know which BCIDs contain protons that will collide. There is also a strong desire to collide as many bunches as possible to maximise the amount of data collected.

2.2 Principles of particle detection

2.2.1 Tracking

Tracking is the general term for determining the path, or *track*, of a particle. A variety of different techniques and technologies can be used but currently they all rely on the particle having an electric charge. As the particle passes through a tracker it will be detected in several positions along its path. These so-called 'hits' can be connected via an algorithm which effectively 'connects the dots' to form the track. The track will be affected by the magnetic field and the momentum of the particle so a precise knowledge of the magnetic field is needed and the particle's momentum will be a free parameter when the track is fitted. More details on this procedure in ATLAS can be found in Section 3.1.

The spatial resolution of the tracker is an important factor in detector design as the precision of the measurement of track position and momentum are dependent on how precisely the position of hits can be determined. This also prevents confusion when there are multiple nearby tracks and a hit could be associated with several different tracks such as the high pileup environment at the LHC. Another important requirement is that any tracking detector is *radiation hard* - resistant to the effects of prolonged exposure to high intensity radiation.

Many modern detectors use semiconductor trackers as they meet these requirements. Whilst the design of specific modules is a very involved process, the general concepts are the same. A semiconductor - often Silicon doped with Germanium - has a bias voltage applied across it which creates a *depletion zone* in the material where there are no mobile charge carriers. When a charged particle passes through this area of so-called 'active' material, it creates electron-hole pairs which are swept by the bias voltage to the cathode and anode respectively. This can be measured as a very small current across the semiconductor junction which is amplified. This amplification can be sensitive to stochastic noise, and experiments often require that the measured current spends a certain Time over Threshold (ToT) to be designated as a hit.

2.2.2 Calorimetry

A calorimeter is a detector intended to measure the energy of particles by stopping them and recording the amount of energy deposited. A high density maximises the interaction cross-section and a large depth ensures that as much of the energy is contained as possible. The energy deposited by a charged particle in different depths of

material is an important factor in calorimeter design. The *stopping power* distribution for muons passing through a copper target is shown in Figure 2.4. Whilst the exact energies and scales depend both on the type of charged particle and the material, the shape and important features remain the same. At lower energies, where ionisation effects dominate, the stopping power decreases as the energy increases until it reaches *minimum ionisation*. At higher energies, the charged particle will lose energy through *bremsstrahlung* radiation and these losses increase rapidly with energy. The stopping power for intermediate energies can be well described by the Bethe formula, also known as the Bethe-Bloch formula [102]

$$-\left\langle \frac{dE}{dx} \right\rangle = \frac{4\pi}{m_e c^2} \cdot \frac{nz^2}{\beta^2} \cdot (\frac{e^2}{4\pi\epsilon_0})^2 \cdot [\ln(\frac{2m_e c^2 \beta^2}{I \cdot (1-\beta^2)}) - \beta^2]. \quad (2.5)$$

For the purposes of this general description it is not necessary to define each term, however one should note that the stopping power is proportional to n, the electron density of the material. This formula only considers the energy loss due to ionisation and so is not valid at higher energies, where radiative effects dominate. When the charged particle radiates, the radiated photon can then further interact with the material, producing a charged particle-antiparticle pair which will themselves radiate more photons. This produces a so-called *shower* of secondary particles each with lower energy than the first. A similar effect occurs when the incident particle interacts with the nuclei of the material, rather than the electrons. These two phenomena are known as electromagnetic and hadronic showers respectively. In many modern particle physics experiments two separate calorimeter systems are used that are each optimised to measure one of these shower types. The typical lengths of the shower formation are known as the radiation length X_0 and the interaction length λ_I. The hadronic interaction length is typically longer than the radiation length so calorimeters looking at hadronic showers will be deeper than those designed for electromagnetic showers. The shape of the shower can be helpful in determining the type of incident particle. In the electromagnetic case, photons will penetrate into the calorimeter further than electrons before beginning their electromagnetic shower. It is therefore useful to segment the calorimeter into layers to allow the measurement of the shower shape.

Decisions about which materials to use when building a calorimeter are driven by two main goals. The calorimeter must be dense enough to absorb the shower quickly and must contain a material sensitive to the ionisation from the secondary particles. Whilst in some cases it is possible to construct a *homogeneous* calorimeter,

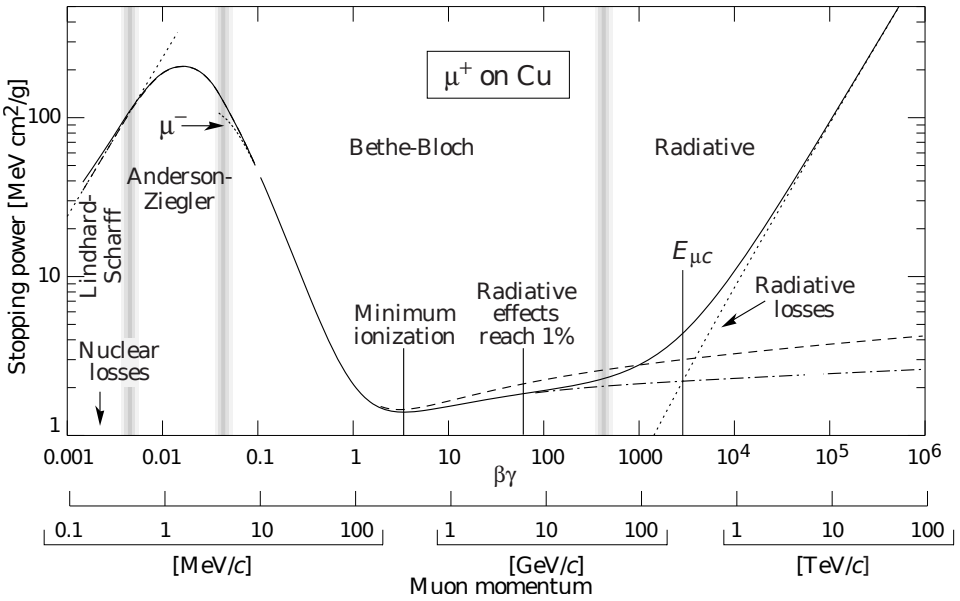


Figure 2.4: Energy deposited by antimuon in copper target. At lower energies, ionisation effects are dominant. At higher energies, the energy losses are mostly due to bremsstrahlung radiation [102].

which uses the same material as the absorber and sampler, they are often very costly. Many modern calorimeters are what is known as *sampling* calorimeters which use layers of different materials to absorb the particles and then detect the secondary shower particles.

The energy resolution of a calorimeter has three main components [103]

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{2.6}$$

where $\oplus$ indicates a quadratic sum. The first term on the right-hand side is called the *stochastic* term and comes about as a result of variations in the shower shape, since the particle shower is a stochastic process. The second term is the *noise* term, which comes from the electronic noise of the calorimeter readout. This can be largely constrained by designing the detector to have a good signal-to-noise ratio. The final term is the *constant* term and comes from nonuniformities in the calorimeter, which could be due to irregular geometries, imperfections in materials or radiation damage, for example. In general, calorimeter resolution improves at higher energies and is worse for low energy showers.

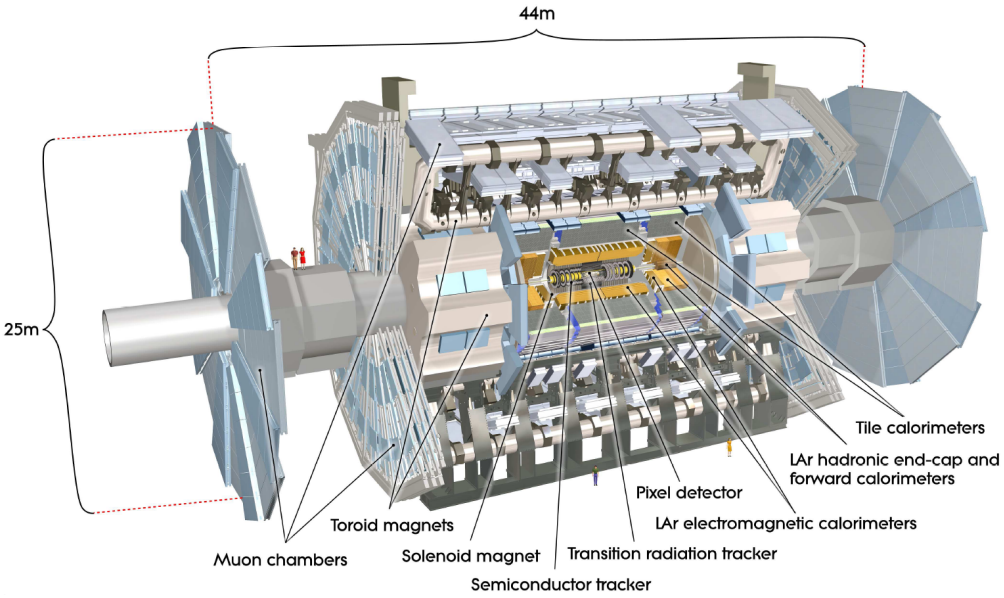


Figure 2.5: Cut-away view of the ATLAS detector [104].

2.3 The ATLAS detector

ATLAS, or *A Toroidal Lhc ApparatuS*, is one of two multi-purpose particle detectors built at the LHC [104, 105]. It has a wide range of physics goals and so has been designed to accurately reconstruct as many different particles as possible. ATLAS makes use of both tracking and calorimetry technology and is composed of many individual sub-detectors which are used in tandem.

2.3.1 Detector Geometry

Before discussing the detector in detail, it will be useful to define the coordinate systems used. A right-handed cartesian coordinate system is sometimes used, with the $+\hat{z}$ direction along the beamline, the $+\hat{x}$ direction is equivalent to the radial direction of the LHC, pointing horizontally away from the centre of the ring and the $+\hat{y}$ direction points vertically upwards.

An alternative set of coordinates is more commonly used, however, which takes advantage of the axial symmetry of the detector. The azimuthal coordinate ϕ is defined to be zero along the vertical direction. Another useful quantity is the rapidity y. It is defined as

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right). \tag{2.7}$$

In the relativistic limit, this can be shown to be equivalent to the *pseudorapidity*

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right), \tag{2.8}$$

where θ is the angle from the beam-axis. It is common to compare positions in the pseudorapidity-azimuthal angular space using the measure of distance ΔR, defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

Sub-detectors are arranged as cylinders around the beamline - known as barrels - with an endcap on each end. This layout maximises the coverage of the detector, so it may be as close as possible to *hermetic*.

2.3.2 Magnets

ATLAS has four magnet systems used to deflect the path of charged particles such that their momenta can be measured. Their layout is shown in Figure 2.6 and their properties are summarised in Table 2.2. The first is a solenoid which surrounds the Inner Detector. It has an inner diameter of 2.46 m, an outer diameter of 2.56 m and a length of 5.8 m and comprises a single coil of Al-stabilised NbTi conductor. It provides a 2 T axial field using the nominal operational current of 7.73 kA.

Table 2.2: Summary of magnet systems in use in the ATLAS detector.

Magnet	**Current / kA**	**Peak field / T**
Barrel solenoid	7.73	2.6 (1.998 axial)
Barrel toroid	20.5	3.9
End-cap toroids	20.5	4.1

A barrel toroid system encloses everything within the calorimeters and consists of 8 separate coils, evenly spaced in ϕ. The construction and materials of each coil are very similar to the solenoid magnet, but the nominal current is higher at 20.5 kA. The peak magnetic field is 3.9 T.

At each end, outside the calorimeter endcaps but within the Muon Spectrometer there is an endcap toroid. Each is effectively a scaled-down copy of the barrel toroid system. The peak field achieved is 4.1 T, slightly higher than for the barrel toroid as the endcap toroids are more compact.

In order to accurately determine the momenta of charged particles, it is necessary to have a precise understanding of the magnetic field throughout the detector. Simulations are used to model the field as the superposition of the Biot-Savart contributions of each magnet system. The magnetisable components within the detector - such as the iron in the support structure - must therefore be well understood.

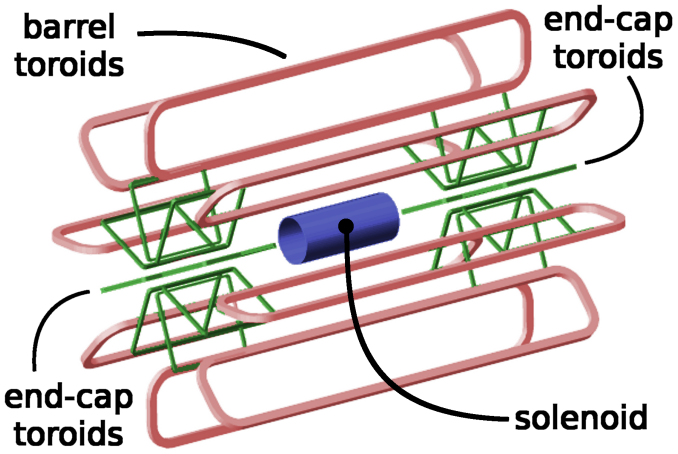


Figure 2.6: Schematic diagram of the ATLAS magnet system [106].

The inner detector is equipped with four NMR probes which can measure the field strength with an accuracy of 0.01 mT. In other areas of the detector, 3D Hall cards - circuit boards with small glass cubes where Hall probes are mounted on each orthogonal face - are used to measure the field strength and direction. These Hall-cards can achieve an accuracy of 0.2 mT for $|\vec{B}| < 1.4$ T and 1 mT for $|\vec{B}| < 2.5$ T.

2.3.3 Inner Detector

The Inner Detector (ID) comprises several sub-detector systems with different underlying technologies but the same goal: detecting charged particles as they are flung from collisions at the beamspot. These sub-detectors are arranged in cylindrical layers as can be seen in Figure 2.7 and the detectors with the finest spatial resolution are placed closest to the beam-line [107]. A summary of these sub-detectors is given in Table 2.3. The combined ability of the Inner Detector to determine the track parameters is highly dependent on the p_{T} and pseudorapidity, but for a typical, central track, the transverse and longitudinal impact parameters can be measured with a resolution of approximately 10 μm and 100 μm respectively and the momentum resolution is approximately 2%.

Pixel detector

The sub-detector placed closest to the beam-line is the Pixel detector, which provides the finest resolution measurements of the passage of charged particles. Each of the 1744 silicon sensors has 47232 pixels and achieves a resolution of 10 μm in the $(R-\phi)$ plane and 115 μm in the z direction. This detector in particular has been

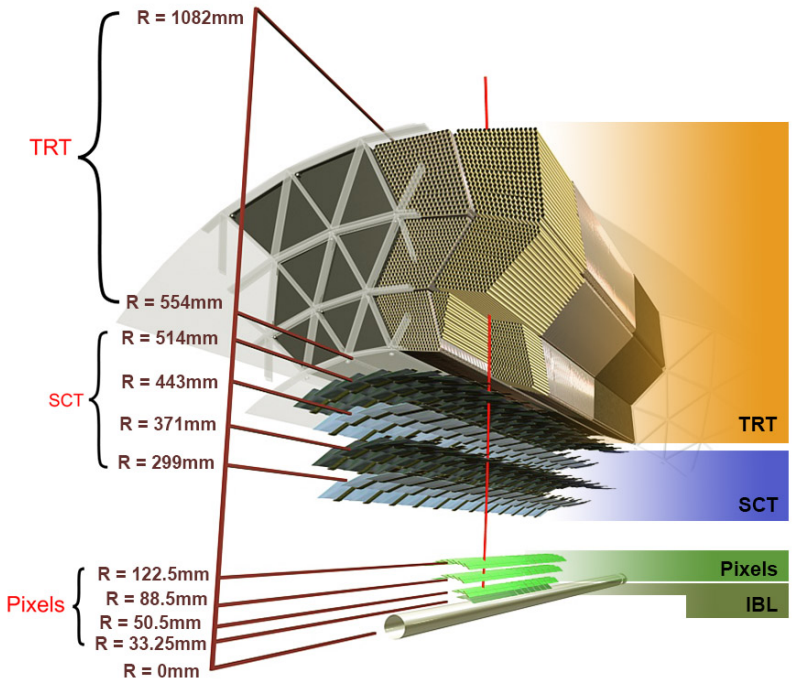


Figure 2.7: A cross-section of the barrel of the ATLAS inner detector [104].

designed to be radiation hard as it is subjected to high particle fluxes. The pixel detector was augmented in May 2014 by the Insertable B-Layer (IBL) [105]. This improved the d_0 measurement resolution by 20% and the $z_0 \sin(\theta)$ measurement by 70%.

Semiconductor Tracker (SCT)

The SCT uses a similar technology and layout to the pixel detector but uses strips of silicon rather than individual pixels. There are four coaxial cylindrical layers in the barrel region and nine disk layers in each end-cap. In total there are 4088 modules which are designed to provide at least four position measurements in the region $|\eta| < 2.5$. Modules are constructed from two p-in-n silicon sensors which are mounted back-to-back. They are rotated by 40 mrad relative to one another to improve the resolution in $(R - \phi)$. The resolution of the SCT is 17 μm in the $(R - \phi)$ plane and 580 μm in the z direction.

Transition Radiation Tracker (TRT)

In order to reduce cost, the outermost tracking detector - the Transition Radiation Tracker (TRT) - uses polyimide drift tubes rather than semiconductor modules.

These 'straws' have a diameter of 4mm and are cut to a length of 144 cm for the barrel and 37 cm for the end-caps. Because of the nature of these straws, the resolution of this detector is the lowest in the ID at 140 μm in the $(R-\phi)$ plane. However, the TRT expects an average of 36 hits per track which allows good pattern recognition for reconstructing the trajectory of passing charged particles.

Table 2.3: Summary of sub-detectors in the ATLAS Inner Detector.

Sub-detector	**Radial extent [mm]**	**η coverage**	**Module resolution [μm]**	
			$(R-\phi)$	z
Pixels	45.5–242	$\lvert\eta\rvert < 2.5$	10	115
SCT	255–610	$\lvert\eta\rvert < 2.5$	17	580
TRT	554–1106	$\lvert\eta\rvert < 2.0$	140	N.A

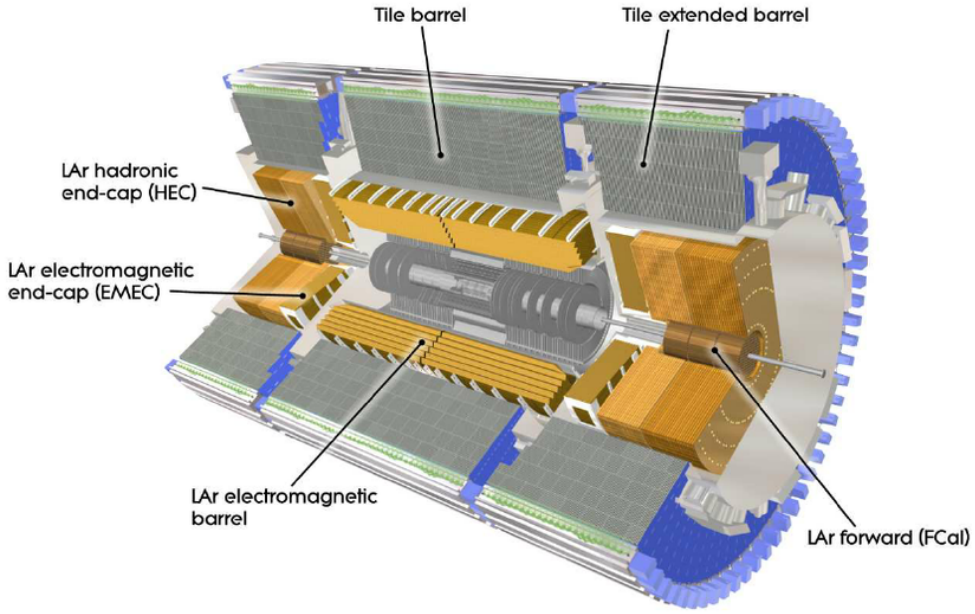


Figure 2.8: Cut-away view of the ATLAS calorimeter system [104].

2.3.4 Calorimeters

ATLAS has two main calorimeter systems, the electromagnetic and hadronic calorimeters, each designed to measure the types of shower described in Section 2.2.2. The layout of the ATLAS calorimeters is shown in Figure 2.8 and a summary of the calorimeter systems is given in Table 2.4.

Electromagnetic Calorimeter

The Electromagnetic Calorimeter is the innermost calorimeter and uses Liquid Argon (LAr) as the sampling material and plates of lead as the absorber [108]. An 'accordian' geometry - where the layers of LAr and lead are laid out in a 'zig-zag' pattern as shown in Figure 2.9 - ensures that the capacitance of the detecting elements is low, and the signal time is short. This layout also results in reasonably uniform performance as a function of ϕ.

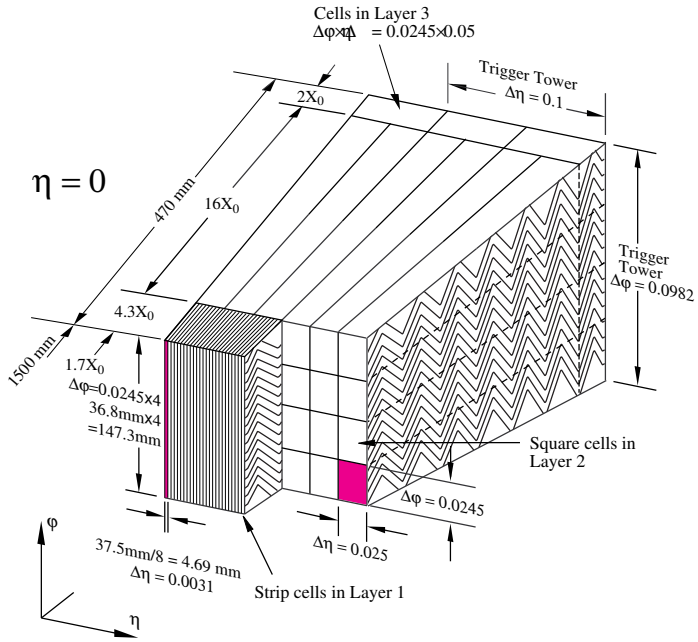


Figure 2.9: Sketch of electromagnetic barrel module showing layout of cells [104].

The modules are segmented radially into a series of layers although the modules of the Electromagnetic Barrel (EMB) - covering a range of $|\eta| < 1.475$ - and the Electromagnetic Endcap (EMEC) - covering $1.375 < |\eta| < 3.2$ are arranged differently. The design of the calorimeter maximises the energy of the electromagnetic shower that is absorbed and the layer segmentation allows the measurement of the showers depth. The first layer has a granularity of $\Delta\eta \times \Delta\phi = 0.0031 \times 0.098$ in the barrel. The second layer, which is designed to absorb most of the shower, is the thickest and spans 16 radiation lengths, X_0. It has a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.0245$ in the barrel. The third layer measures any remaining energy in the tail of the shower and needs less granularity in pseudorapidity - $\Delta\eta \times \Delta\phi = 0.05 \times 0.0245$ in the barrel.

The energy lost in the ID, magnets and services before reaching the calorimeter must be taken into account and for this purpose an 11 mm thin, active LAr layer

is placed just within the electromagnetic calorimeter which is used to calibrate the energy measurement. This layer is known as the *presampler*.

After the electromagnetic calorimeter has been calibrated it provides an energy measurement with a resolution of $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E(\text{GeV})}} \oplus 0.7\%$ [109].

Hadronic and forward calorimeters

The hadronic calorimeter is actually composed of three separate calorimeters: the tile calorimeter [110] - which uses a different technology to the electromagnetic calorimeter - the liquid-argon Hadronic Endcap Calorimeter (HEC) and Forward Calorimeter (FCal).

The tile calorimeter uses steel tiles as the absorber as such a large quantity of lead would be too heavy to support without compromising performance. It also uses a plastic scintillator as the active medium; the scintillation light is collected by wavelength-shifting fibres which are connected to Photomultiplier Tubes (PMTs). It covers a region of $|\eta| < 1.7$ and is subdivided into the central barrel, 5.8 m in length and two extended barrels each 2.6 m in length. It extends from a radius of 2.28 m to 4.25 m from the beamline and corresponds to approximately 7.4λ (hadronic interaction lengths). Therefore approximately 1 in 2000 hadronic showers will *punch through* to the muon spectrometer which must be taken into account. The tile calorimeter has a granularity of $\Delta\eta \times \Delta\phi \approx 0.1 \times 0.1$ for the innermost 2 layers. Cells in the outer layer are twice as large in the η direction.

Table 2.4: Summary of ATLAS calorimeters.

Sub-detector	**Radial extent [m]**	**η coverage**	**Resolution**
LAr	1.5–1.97	$\lvert\eta\rvert < 3.2$	$\frac{10\%}{\sqrt{E(\text{GeV})}} \oplus 0.7\%$
Tile	2.28–4.25	$\lvert\eta\rvert < 1.7$	$\frac{50\%}{\sqrt{E(\text{GeV})}} \oplus 3\%$
HEC	0.37–2.03	$1.5 < \lvert\eta\rvert < 3.2$	$\frac{50\%}{\sqrt{E(\text{GeV})}} \oplus 3\%$
FCal	< 0.44	$3.1 < \lvert\eta\rvert < 4.9$	$\frac{100\%}{\sqrt{E(\text{GeV})}} \oplus 10\%$

The HEC and FCal which occupy the forward regions ($1.5 < |\eta| < 3.2$ and $3.1 < |\eta| < 4.9$ respectively) are exposed to higher intensity irradiation and so LAr is used as the argon can be easily replaced to prevent radiation damage. For the HEC, copper plates are used as the absorber. In the FCal copper plates are also

used on the inner layers to measure electromagnetic showers whereas tungsten is used in the outer layers to measure hadronic showers.

The energy resolution of the hadronic calorimeter in the barrel and endcap is approximately $\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E(\text{GeV})}} \oplus 3\%$ [111,112]. In the case of the FCal, the energy resolution is approximately $\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E(\text{GeV})}} \oplus 10\%$ [113].

2.3.5 Muon Spectrometer

The Muon Spectrometer (MS) forms the outermost part of the ATLAS detector and detects any charged particles that manage to pass through the calorimeters. It has two primary objectives: to provide momentum measurements in the region $|\eta| < 2.7$ and to trigger on particles within $|\eta| < 2.4$.

There are four types of module used in the MS, each using different technologies. As shown in Figure 2.10, there are Monitored Drift Tubes (MDTs), Cathode-Strip Chambers (CSCs), Thin Gap Chambers (TGCs) and Resistive Plate Chambers (RPCs). The former two are used to accurately measure particle momenta and the latter two are used for triggering.

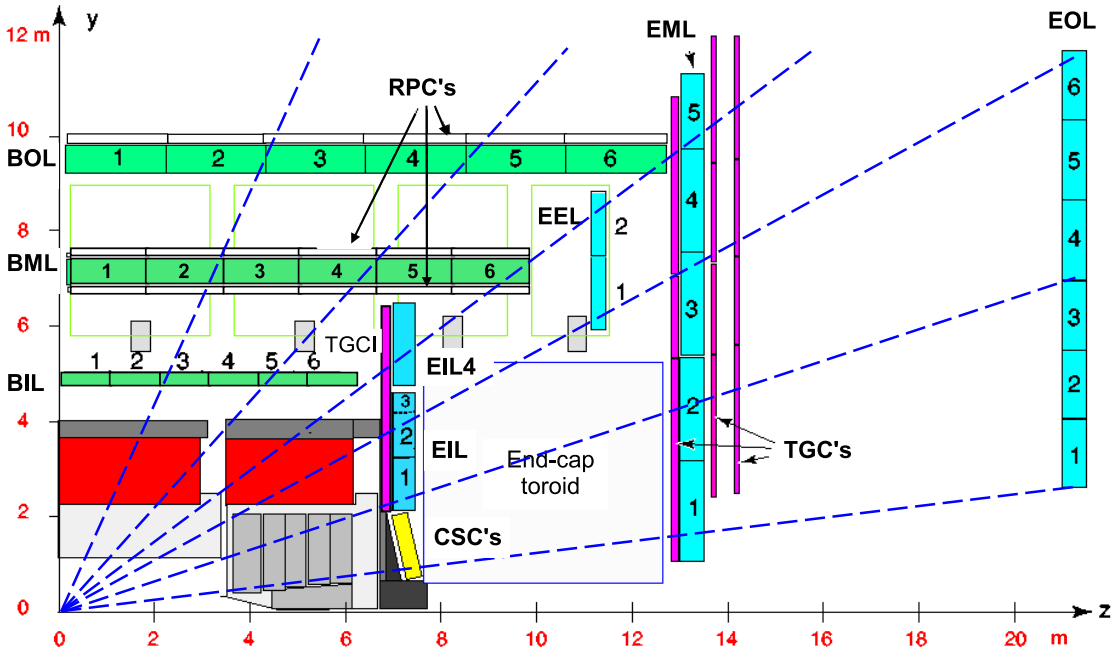


Figure 2.10: Cross-section of the ATLAS Muon Spectrometer showing the MDT modules in green and blue as well as the positions of the CSCs, RPCs and TGCs [104].

An MDT is a drift chamber composed of aluminium tubes filled with a mixture of argon and CO_2 and a central wire constructed from a tungsten-rhenium alloy which collects electrons liberated from the gas by high-energy charged particles. The CSCs modules use similar materials but are able to safely operate at higher counting rates so they are used in the forward region $|\eta| > 2$. They are *multi-wire proportional* chambers, with many parallel anode wires placed between cathode strips as shown in Figure 2.11. The charges on each wire are not read out, instead the position of the track is found by interpolating the charge on adjacent cathode strips.

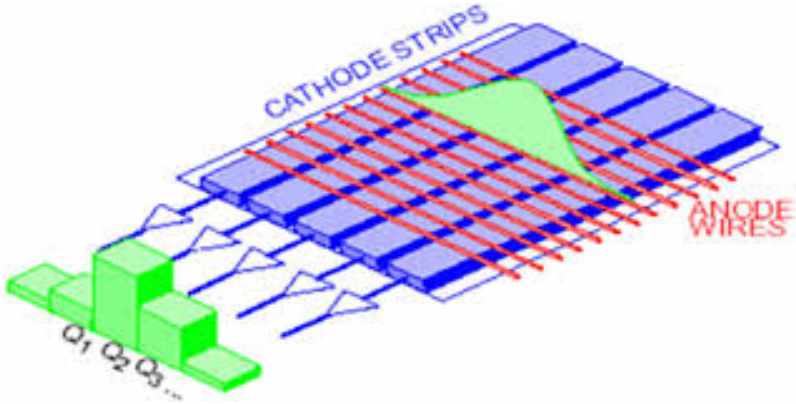


Figure 2.11: Schematic of the layout of a Cathode-Strip Chamber (CSC) module. The anode wires are placed between cathode strips and the position of the track is interpolated from the charges on adjacent strips, shown in green [104].

The MDT and CSC 'precision modules' are complemented by modules that have better time resolution and are able to provide the fast readout needed for triggering. In the barrel there are three layers of RPCs, which use parallel plates made from phenolic-melaminic plastic laminate separated by 2 mm and kept at a potential difference of 9.8 kV. The endcap contains TGCs which are multi-wire proportional chambers, similar to the CSCs but operating at a higher voltage and with a smaller gap between the wires. This reduces the drift time and improves time resolution.

The path of charged particles is strongly affected by the magnetic field generated by the toroid magnets which ranges in magnitude from 0.15–2.5 T. The momentum resolution is given by $\frac{\delta p}{p} = \Delta S \times \frac{p(\text{TeV})}{500\ \mu\text{m}}$ where ΔS is the sagitta resolution which for a muon crossing three MDT chambers is approximately 45 μm [114].

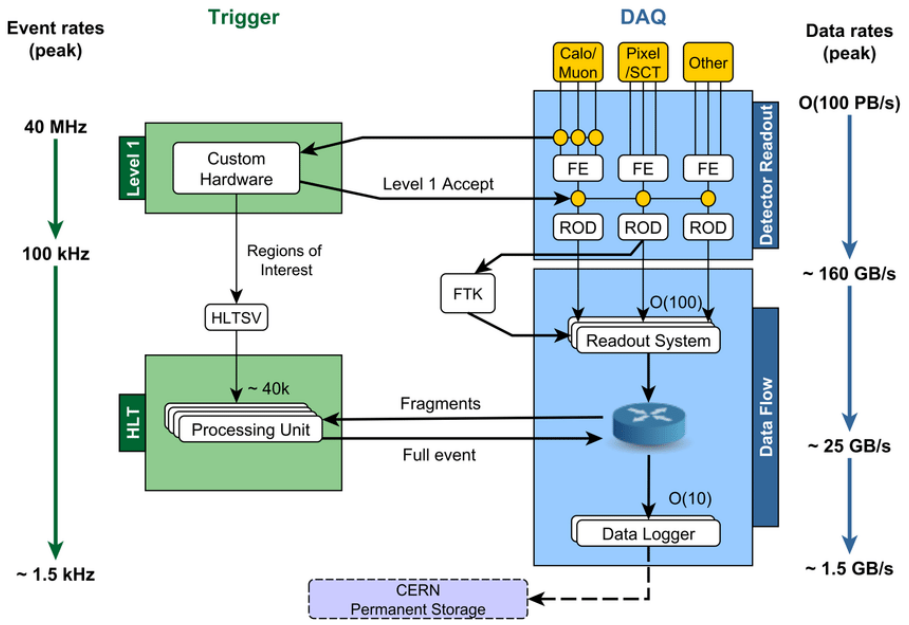


Figure 2.12: Diagram of the ATLAS Trigger and Data Acquisition (TDAQ) systems. The peak event and data rates at each stage are shown on the left and right respectively [116].

2.4 The ATLAS Trigger system

With bunches colliding at a rate of 40 MHz, it is not feasible to read out every event from the detector and write this to disk. Furthermore, most of these events will be uninteresting. The purpose of the ATLAS trigger system is to identify those events that are of interest so that they can be recorded at a reasonable rate. This section will describe the organisation of the ATLAS trigger in general. Later, in Section 3.5.1, a more detailed discussion of the muon triggers used in the analysis will be given.

The trigger comprises two levels: the Level 1 (L1) trigger uses only hardware to reduce the acceptance rate to approximately 100 kHz. The High-Level Trigger (HLT) uses software to reconstruct events more completely and aims to reduce the acceptance rate to approximately 1.4 kHz [115]. Since ATLAS is aiming to collect events with many different physics signatures, a *menu* is created which contains algorithms arranged into *chains* each corresponding to a particular signature. Some objects such as jets are fairly ubiquitous in ATLAS and so it may be necessary to reduce the rate at which ATLAS triggers on these objects. This is achieved by using a prescale, n, such that the relevant triggers will only accept $\frac{1}{n}$ events.

2.4.1 Level 1 Trigger

With a bunch crossing every 25 ns, there is little time to make a trigger decision. The first stage of the trigger, L1, is solely hardware-based and is designed to make an accept/reject decision in under 2.5 μs. Only certain sub-detectors are able to be read out at such a fast rate and the L1 system comprises several components designed to collate reduced-granularity data from different sub-detectors.

The L1Calo system reads out from the calorimeters. A full reconstruction of energy deposits in the calorimeter would be too slow to be useful and so a sliding window fit is performed, searching for local maxima in so-called *trigger towers*. Trigger towers are formed from the analogue sums of calorimeter cells in regions of coarse granularity in $\Delta\eta \times \Delta\phi$ from 0.1×0.1 in the central region to 0.4×0.4 in the forward regions. Electron and Photon L1 candidates are defined as regions of 2×2 trigger towers in the EM calorimeter. Jet candidates, including hadronically-decaying taus are built from trigger towers in both EM and hadronic calorimeters.

The L1Muon system reads out from the RPC and TGC modules of the Muon Spectrometer. A recent addition to the L1 trigger is L1Topo which can combine information from the L1Muon and L1Calo and calculate basic topological quantities such as the invariant mass of candidate objects. The overall L1 accept decision is made by the Central Trigger Processor (CTP) which checks if the bunch crossing contains an event that passes the requirements of an item in the trigger menu. The CTP sends the Region of Interests (RoIs) of the trigger objects as seeds for the HLT selection.

2.4.2 High-Level Trigger

The HLT attempts to match the offline event reconstruction described in Section 3 as closely as possible. The exact algorithms that are run will depend on which L1 items have passed. Algorithms are arranged into chains which are each seeded by at least one L1 item. Whilst running, ATLAS has several data *streams* each with a different purpose. For instance, there are streams for calibration, debugging and of course, physics analysis. Each HLT chain can be used to trigger on events for multiple streams. The HLT is designed to exhibit *early rejection* - at several stages throughout a chain's execution there are 'hypothesis' algorithms which decide if the trigger is still likely to accept and therefore whether this chain should continue to be executed. A given event can be rejected if there are no remaining chains active, freeing up CPU resources for other events. The HLT runs on a processing farm which is run on rack-mountable server-class PCs, connected by multi-layer gigabit-

Ethernet networks [116]. If an event passes an HLT chain, the entire detector readout is serialised and sent to the Tier-0 CERN Data Centre where it will be written to disk to be reconstructed offline [117].

2.5 Trigger Automated Cost Monitoring

In order to maximise the amount of data recorded, it is necessary to push the limited computing resources as far as possible. The HLT runs on a CPU farm with approximately 50,000 cores. Even so, CPU usage must still be closely monitored and in particular, the time taken to process each event must be strictly controlled. Another important consideration is the rate at which the readout electronics are running. If the trigger system requests data at too high a rate for too long, it could cause damage to the detector. The task of 'cost-monitoring' - checking the trigger's usage of resources - is important to its smooth operation and minimising downtime when taking data.

As the HLT is running it can store information about the resources it used and the execution times of its algorithms. By default, this 'cost mode' is used only in 1 out of 10 events in order to prevent the cost-monitoring from becoming a large drain on resources itself. The stored cost information will be written out along with the rest of the event data. In the past, this data was analysed manually to check for any problems, which was time consuming and prone to human error. As part of this author's contribution to the ATLAS collaboration, a software framework was developed to analyse the cost-monitoring data automatically and produce alerts if the usage exceeds a designated threshold.

It is useful to look at the HLT execution with different levels of granularity, for instance looking at how an entire chain executes as well as each individual algorithm within that chain. For this purpose a 'Monitor' object is created for each chain, algorithm etc. and each Monitor contains many 'Counters' which record the interesting variables (e.g execution time) that should be analysed later. This level of granularity is very useful in tracking down problems in the trigger, but it is time-consuming to check every Counter of every Monitor, hence the desire to automate this procedure.

The Automated Cost Monitoring (AutoCostMon) framework was developed to allow trigger users to easily specify which cost-monitoring variables they want to check and the thresholds for when an alert should be created. The general structure of the framework is shown in Figure 2.13. The configuration of the AutoCostMon framework can be controlled from a single `XML` file. Currently, there is a web-display,

shown in Figure 2.14, which shows the important variables and allows ATLAS users to see how the trigger is performing. It is intended that the results of the automated cost monitoring will be integrated into this web display, for instance adding red warning icons to items that exceed their expected values.

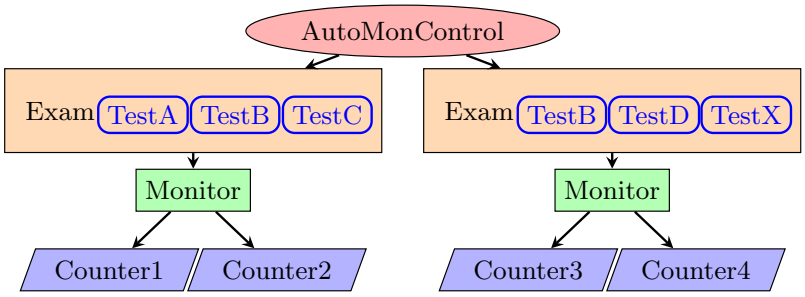
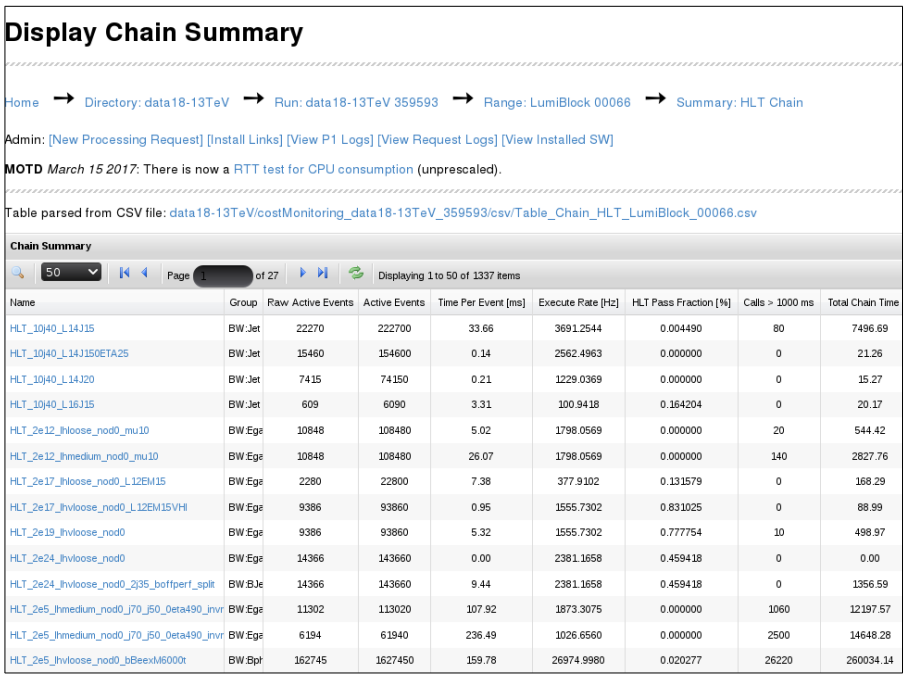


Figure 2.13: Organisational structure of Automated Cost Monitoring. An 'exam' is a series of tests that can be run on a monitor, for instance an HLT chain, and checks the values of the contained counters. This structure can be completely configured in a single `XML` file.

During LHC Run 2, there was considerable effort to improve the CPU efficiency of the HLT. The existing ATLAS cost-monitoring framework consumed approximately 2.5% of the total HLT CPU time which, although only a small fraction, was a target for optimisation. A large number of minor bugs were fixed and changes to control flow were made to ameliorate the CPU usage. The most effective change was the replacement of ordered `STL` maps, used for caching HLT data, with unordered maps that have far more efficient lookup times (amortized $\mathcal{O}(1)$ vs. $\mathcal{O}(log\ n)$). This simple change significantly reduced the CPU usage of the cost-monitoring framework, as shown in Figure 2.15.

2.6 Performance of ATLAS during Run 2

The end of 2018 marked the end of LHC Run 2, and has been followed by Long Shutdown 2 (LS2) where a number of upgrades to the accelerator and experiments are planned. The total integrated luminosity over the course of Run 2 is shown in Figure 2.16. One should note that the rate of data-collection was much higher towards the end Run 2, reflecting how optimisations made - and experience gained - over this time helped to maximise the performance of the accelerator and experiment. The dataset studied in this thesis contains the data collected in 2015 and 2016 and corresponds to approximately 36 fb^{-1} of integrated luminosity.



Display Chain Summary

Home → Directory: data18-13TeV → Run: data18-13TeV 359593 → Range: LumiBlock 00066 → Summary: HLT Chain

Admin: [New Processing Request] [Install Links] [View P1 Logs] [View Request Logs] [View Installed SW]

MOTD *March 15 2017*: There is now a RTT test for CPU consumption (unprescaled).

Table parsed from CSV file: data18-13TeV/costMonitoring_data18-13TeV_359593/csv/Table_Chain_HLT_LumiBlock_00066.csv

Chain Summary

50 | Page 1 of 27 | Displaying 1 to 50 of 1337 items

Name	Group	Raw Active Events	Active Events	Time Per Event [ms]	Execute Rate [Hz]	HLT Pass Fraction [%]	Calls > 1000 ms	Total Chain Time
HLT_10j40_L14J15	BW:Jet	22270	222700	33.66	3691.2544	0.004490	80	7496.69
HLT_10j40_L14J150ETA25	BW:Jet	15460	154600	0.14	2562.4963	0.000000	0	21.26
HLT_10j40_L14J20	BW:Jet	7415	74150	0.21	1229.0369	0.000000	0	15.27
HLT_10j40_L16J15	BW:Jet	609	6090	3.31	100.9418	0.164204	0	20.17
HLT_2e12_lhloose_nod0_mu10	BW:Ega	10848	108480	5.02	1798.0569	0.000000	20	544.42
HLT_2e12_lhmedium_nod0_mu10	BW:Ega	10848	108480	26.07	1798.0569	0.000000	140	2827.76
HLT_2e17_lhloose_nod0_L12EM15	BW:Ega	2280	22800	7.38	377.9102	0.131579	0	168.29
HLT_2e17_lhvloose_nod0_L12EM15VHI	BW:Ega	9386	93860	0.95	1555.7302	0.831025	0	88.99
HLT_2e19_lhvloose_nod0	BW:Ega	9386	93860	5.32	1555.7302	0.777754	10	498.97
HLT_2e24_lhvloose_nod0	BW:Ega	14366	143660	0.00	2381.1658	0.459418	0	0.00
HLT_2e24_lhvloose_nod0_2j35_boffperf_split	BW:BJe	14366	143660	9.44	2381.1658	0.459418	0	1356.59
HLT_2e5_lhmedium_nod0_j70_j50_0eta490_invr	BW:Ega	11302	113020	107.92	1873.3075	0.000000	1060	12197.57
HLT_2e5_lhmedium_nod0_j70_j50_0eta490_invr	BW:Ega	6194	61940	236.49	1026.6560	0.000000	2500	14648.28
HLT_2e5_lhvloose_nod0_bBeexM6000t	BW:Bph	162745	1627450	159.78	26974.9980	0.020277	26220	260034.14

Figure 2.14: Screenshot of web-display of ATLAS Trigger cost-monitoring data

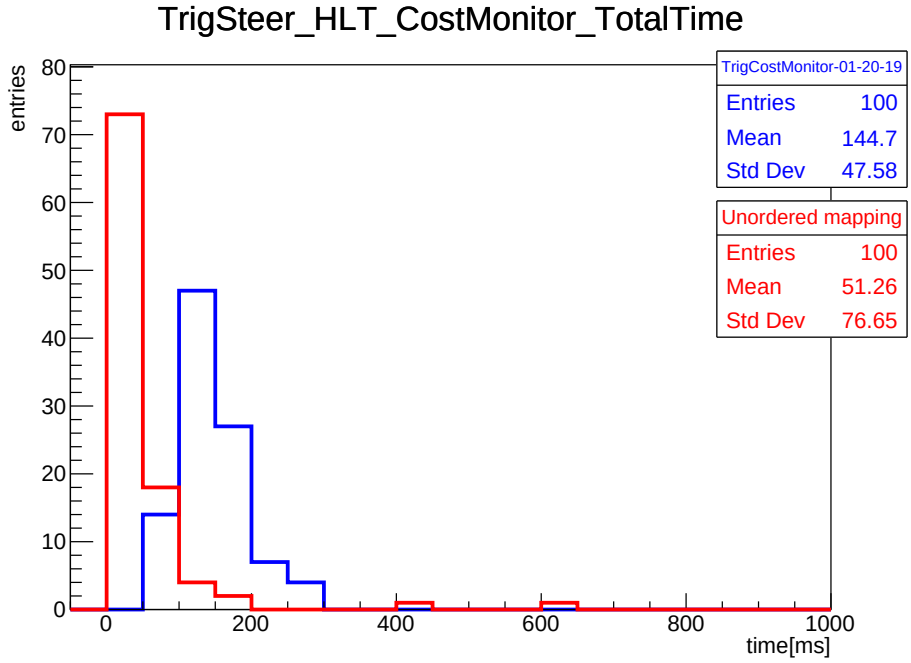


Figure 2.15: Cost monitoring execution time using ordered maps (blue) compared to unordered maps (red).

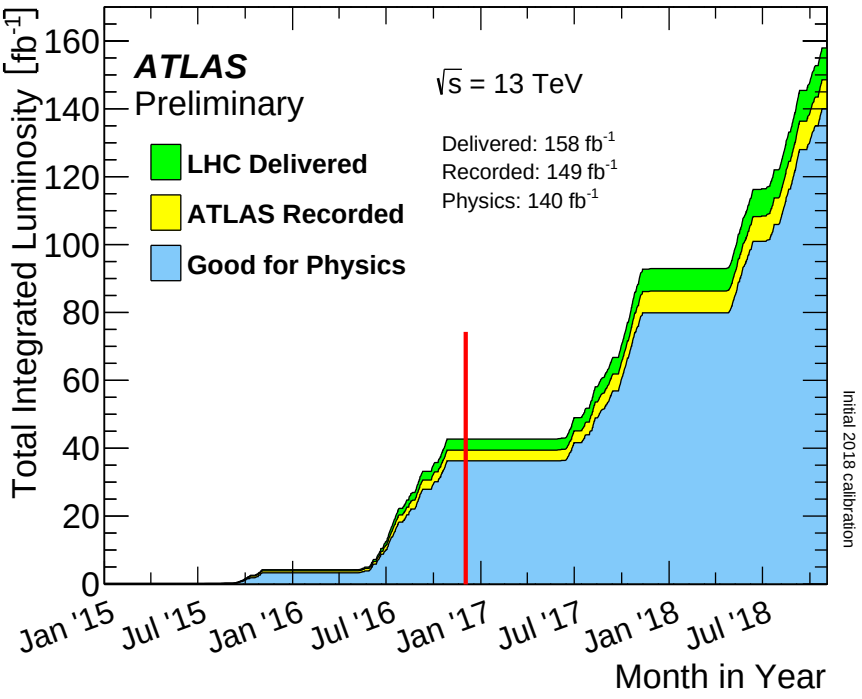


Figure 2.16: The total integrated luminosity during Run 2 that was delivered by the LHC (green), recorded by ATLAS (yellow) and deemed to be high enough quality for physics analysis (blue). The dataset analysed in this thesis corresponds to data collected in 2015 & 2016, left of the red line [99].

Chapter 3

Object Reconstruction

What is real? How do you define 'real'? ... 'real' is simply electrical signals interpreted by your brain.

Morpheus

When an event has been selected by the trigger, the raw readout of the detector is stored to be reconstructed later. This reconstruction comprises a series of many algorithms each designed to combine signals from different sub-detectors and interpret which physics objects have been detected. The first step is to identify the tracks of charged particles and clusters of deposited energy in the calorimeters, which are then used by particle reconstruction and identification algorithms.

3.1 Tracking and Vertexing

As described in Section 2.2.1, charged particles passing through the Inner Detector leave hits at various points along their path. There are many different approaches to reconstructing a particles track; the primary reconstruction algorithm in ATLAS is an 'inside-out' procedure that consists of the following steps [118]:

– **Clustering:** The ionisation charge from a passing particle can often be swept to more than one Pixel or Semi-conductor tracker (SCT) cell. Clusters are formed by finding these connected cells but in some cases nearby particles can be clustered together as shown in Figure 3.1. To ameliorate this, a neural network based algorithm is employed to identify these merged clusters [119].

– **Formation of space-points:** Space-points are three dimensional measurements - and uncertainties - of a charged particle's position as it traversed the active material of the Inner Detector. In the Pixel detector, one pixel cluster equates to one space-point whereas in the SCT, clusters from both sides of a strip layer are combined to form a space-point. In either case, it is necessary to transform from a position relative to the module to the ATLAS global coordinate system.

– **Track finding:** Initial track seeds are formed from 3 space-points, this allows crude estimates of the momentum and impact parameters. A combinatorial Kalman filter [120] is used to build track candidates from the seeds by incorporating additional space-points in the Pixel and SCT detectors which are compatible with the preliminary trajectory. Multiple track candidates can be constructed for the same seed if there are multiple compatible space-point extensions on the same layer.

– **Ambiguity solving:** When space-points are shared, the ambiguity over which track candidate the space-point corresponds to must be resolved. A track score is calculated from several track parameters, including the number of space-points in each detector, the χ^2 goodness-of-fit and the *holes* in each detector. A hole is a layer of sensitive material in the track's path where one would expect to see a cluster but no such cluster exists. Track candidates are ranked by their track score, and are iteratively processed by removing shared space-points until all track candidates have either passed or failed a set of simple quality requirements [119].

– **TRT extension:** Track candidates within the coverage of the TRT are extended into the TRT where compatible space-points are added to the track. Space-points in the TRT are not constructed from clusters, as in the Pixel and SCT but rather *drift circles*, which are calculated from the time taken for the ionisation charge to drift to the high-voltage wire in each TRT straw.

– **Final Fit:** A final fit to all of the space-points in a track is performed. The addition of TRT space-points greatly improves the momentum resolution [121].

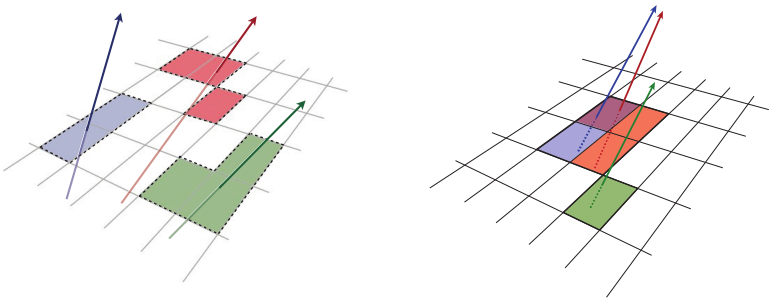


Figure 3.1: Typical single-particle pixel clusters (left). When particle trajectories become too close they are clustered together, forming a single merged cluster (right). A neural network based algorithm is used to identify merged clusters and attempts to recover suitable single-particle clusters [119].

Muons pass through the entire detector, leaving a track in the Inner Detector but also in the Muon Spectrometer (MS). Tracks in the MS are formed through a similar procedure to that described above for the Inner Detector, however in this case tracks are seeded from the outermost layers and the track is extrapolated inwards [122]. If a track in the MS is found to be compatible with a track in the Inner Detector, it is considered a *combined* track and a refit of the track parameters is performed, using measurements from both detectors. Muons with combined tracks benefit from better momentum and impact-parameter resolution.

Vertexing

In addition to the identification of the tracks themselves, the vertex from which the track originates is also of interest. The *beamspot*, the area where the majority of collisions take place, is only approximately 14 μm in the transverse direction but approximately 45 mm in the longitudinal direction [123] as shown in Figure 3.2. Therefore, vertices are best distinguished by their longitudinal position along the beamline. The z_0-position of tracks are used in an iterative χ^2 fit as a vertex seed. If another track lies within seven standard deviations of this seed, it is added to the vertex, otherwise it will seed a new vertex. This procedure continues until no more interaction vertices can be found. For hard scatter processes, the primary vertex is defined to be the vertex with the largest sum of squared transverse momenta, $\sum p_T^2$, where p_T are the transverse momenta of the tracks associated with the vertex. This assumes that this vertex is the location of the largest momentum exchange, where particles with large masses are most likely to be created [124]. The impact parameters of other tracks, for instance the tracks associated to b-jets, are calculated relative to this primary vertex.

Other vertices belong to two categories: The first are *pile-up* vertices, which are caused by other proton-proton interactions than that of the primary vertex. A commonly-used measurement of pile-up is the mean number of interactions per bunch crossing, $< \mu >$, and in Figure 3.3 one can see that a typical bunch crossing in ATLAS results in between roughly 10 and 60 interactions.

If one of the decay products from the primary vertex itself decays after a non-negligible time, this will result in a *secondary* vertex. These are of particular interest when searching for b-quarks, as the lifetime of b-hadrons is typically long enough, $\mathcal{O}(1.5 \times 10^{-12}$ s), that they decay a significant distance from their creation at the primary vertex $\mathcal{O}(1$ mm) [8].

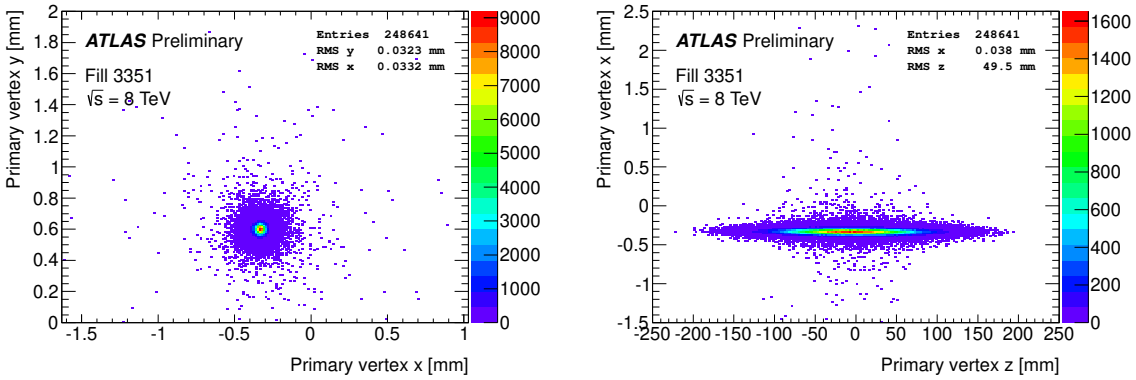


Figure 3.2: The ATLAS beamspot is symmetric about the beamline with a transverse width much smaller than the length. During Run 1 of the LHC, the beamspot was larger than in Run 2, with dimensions of approximately $32\mu m \times 32\mu m \times 50$mm compared to approximately $14\mu m \times 14\mu m \times 45$mm [123].

3.2 Topological Clustering

Showers in the calorimeters will deposit energy in many connected cells. So, just as in the case of the Pixel and SCT clusters used for tracking, it is useful to construct *topological clusters* of energy deposits in the calorimeters. The cell signal significance, ζ_{cell}, is related to the energy deposited in the cell, E_{cell} and the expected average noise, $\sigma_{\text{noise, cell}}$ by [125]:

$$\zeta_{\text{cell}}^{\text{EM or LCW}} = \frac{E_{\text{cell}}^{\text{EM or LCW}}}{\sigma_{\text{noise, cell}}^{\text{EM or LCW}}}, \tag{3.1}$$

where the measurement can be made at the Electromagnetic Scale (EM) or using Local Hadronic Cell Weighting (LCW). The EM calibration scheme was determined

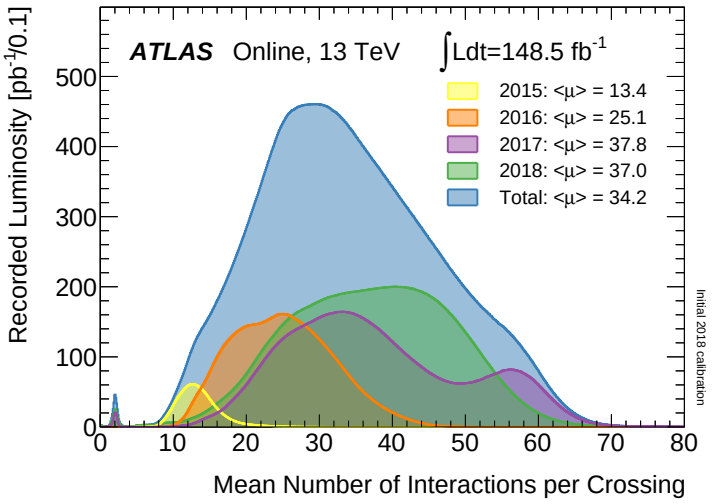


Figure 3.3: The mean number of interactions per bunch crossing, $< \mu >$, for LHC Run 2. In later years (2017 & 2018), ATLAS operated at higher pile-up than the earlier years, to maximise the integrated luminosity collected [99].

by measurements made in an electron test beam and correctly reconstructs the energy deposits from electrons and photons. The LCW scheme calibrates each cluster based on additional information such as the depth of the cluster in the calorimeter.

The growth of a cluster is controlled by three parameters $\{S, N, P\}$, which are the primary seed threshold, the threshold for growth control, and the principal cell filter respectively. A cell with significance $|\zeta_{cell}| > S$ forms a cluster seed. Neighbouring cells - either directly adjacent in the same layer, or in adjacent layers and overlapping in (η, ϕ) - are added to the cluster if they pass the threshold $|\zeta_{cell}| > P$. If an added cell passes the threshold $|\zeta_{cell}| > N$, then its neighbours are also added to the cluster. If a seed-cell of another cluster is added, the two clusters are merged into a single cluster. The default parameters used in ATLAS are $\{S, N, P\} = \{4, 2, 0\}$.

This procedure works well for single, isolated particle showers but in the high pile-up environment of normal running conditions at ATLAS, it is common for these proto-clusters to contain multiple showers. Therefore, a cluster-splitting algorithm is used which splits clusters containing two or more local-maxima (cells with deposited energy $E_{cell} > 500$ MeV). Cells which neighbour two local maxima can be shared, contributing a fraction of their energy to each new cluster. The exact fraction is calculated from a weight determined by the energies of the two maxima, as well as their distance to the cell in question [125].

3.3 Jets

Jets are collimated 'sprays' of particles that are the result of fragmentation of high-energy gluons and quarks. They are particularly prevalent at hadron colliders, like the LHC, where a major challenge for accurate event reconstruction is identifying which particles should be associated with the jet, and which are from pile-up interactions. There are a plethora of algorithms for this purpose, each with their own advantages and drawbacks but they can be broadly separated into two categories [126]. *Cone algorithms* assume that particles occupy conical regions in the detector and cluster all objects within a certain radius, thus the jets have a fixed shape and size. Whilst easy to implement, these algorithms have fallen out of favour as they are not Infra-red and Collinear (IRC) safe [127]. High-energy partons will radiate lower-energy gluons as part of the non-perturbative fragmentation process. In order to accurately compare theoretical predictions with experimental measurements, the contributions of this radiation to the energy deposited must be consistently assigned to a jet. A jet-finding algorithm that is unaffected by the addition of gluons either in the low-energy or small-angle limit is said to be IRC safe.

Most modern particle physics experiments use *sequential algorithms*, where jets are clustered by adding objects one-by-one if they pass a threshold. These algorithms often start by defining two distance measurements: d_{ij} is the distance between two objects (tracks, topo-clusters etc.) and d_{iB} is the distance of an object to the beam:

$$d_{ij} = min(p_{Ti}^{2n}, p_{Tj}^{2n})\frac{\Delta_{ij}^2}{R^2}, \tag{3.2}$$

$$d_{iB} = p_{Ti}^{2n}, \tag{3.3}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and p_{Ti}, y_i and ϕ_i are the transverse momentum, rapidity and azimuthal angle of object i. The radius parameter R is selected as a compromise; large-R jets are more likely to contain all of the jet energy but are susceptible to also including energy from pile-up interactions. Usually a value between 0.4 and 0.7 is chosen. These algorithms proceed by finding the set $\{d_{ij}, d_{iB}\}$ of distances for all objects, i, j. If the distance d_{iB} is the minimum, object i is defined as a final jet and removed from the list of objects. However, if for some object j, the distance d_{ij} is the minimum, it is combined with object i by summing their four-momenta and this new object is returned to the list of objects so the process can begin again. This procedure repeats iteratively until every object has been assigned to a jet. Historically, different choices of the parameter n have been used

which affects the shape of the resulting jets as seen in Figure 3.4. ATLAS uses $n = -1$ which defines the IRC safe *anti-k_t* algorithm, which has been shown to be particularly insensitive to the effects of pile-up [127].

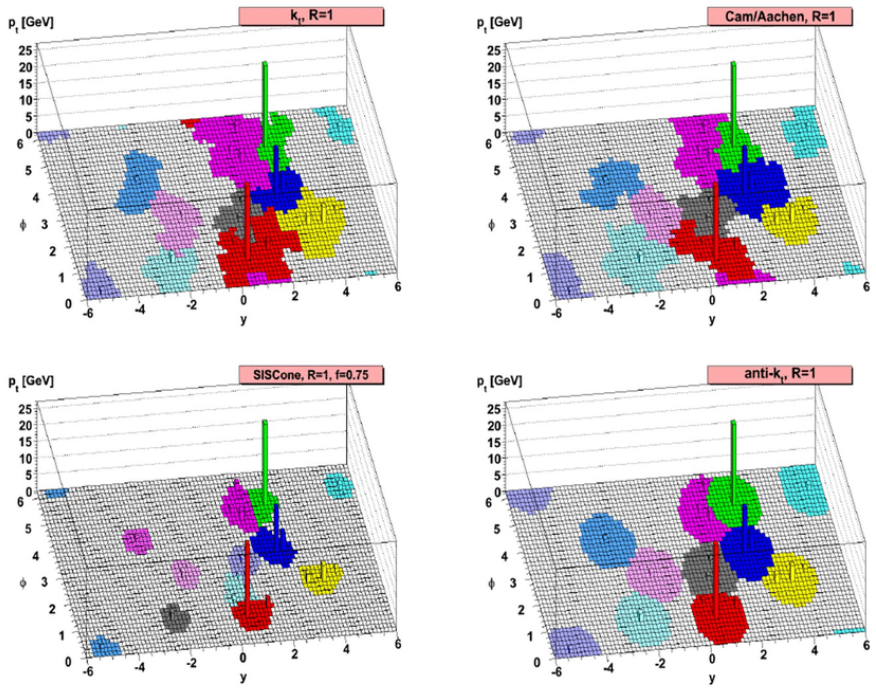


Figure 3.4: The result of using different jet clustering algorithms on the same energy deposits in a calorimeter. Different jet clustering algorithms produce jets with markedly different shapes. The anti-k_t algorithm (bottom right) produces jets that are almost conical, and has been shown to be particularly insensitive to the effects of pile-up [127, 128].

3.4 Flavour Tagged Jets

Decay modes involving heavier quarks such as the charm and bottom are often particularly interesting to physicists. For instance, the dominant decay channel of the Standard Model Higgs is $H \to b\bar{b}$. It is therefore useful to identify if a jet originated from one of these flavours of quark. A useful property of b-quarks is that they form b-hadrons which have a non-negligible lifetime. This means that they will travel away from the interaction point before decaying, leaving a secondary vertex. The tracks from the b-hadron decay products will therefore have large impact parameters with respect to the primary vertex.

There are three widely-used methods for searching for b-jets [129]:

– **Impact parameter:** The IP2D algorithm uses the significance of the *transverse* impact parameter, $\frac{d_0}{\sigma_{d_0}}$ where σ_{d_0} is the uncertainty on the reconstructed impact parameter. The IP3D algorithm also uses the transverse impact parameter, but in addition uses the significance of the *longitudinal* impact parameter $\frac{z_0}{\sigma_{z_0}}$ where again, σ_{z_0} is the uncertainty. These algorithms both use a Log-Likelihood Ratio (LLR) method where the impact parameter significances are compared to probability-density functions for the b-jet and light-jet hypotheses and the ratio of these probabilities defines the track weight. The total jet weight is the sum of the logarithms of individual track weights [130].

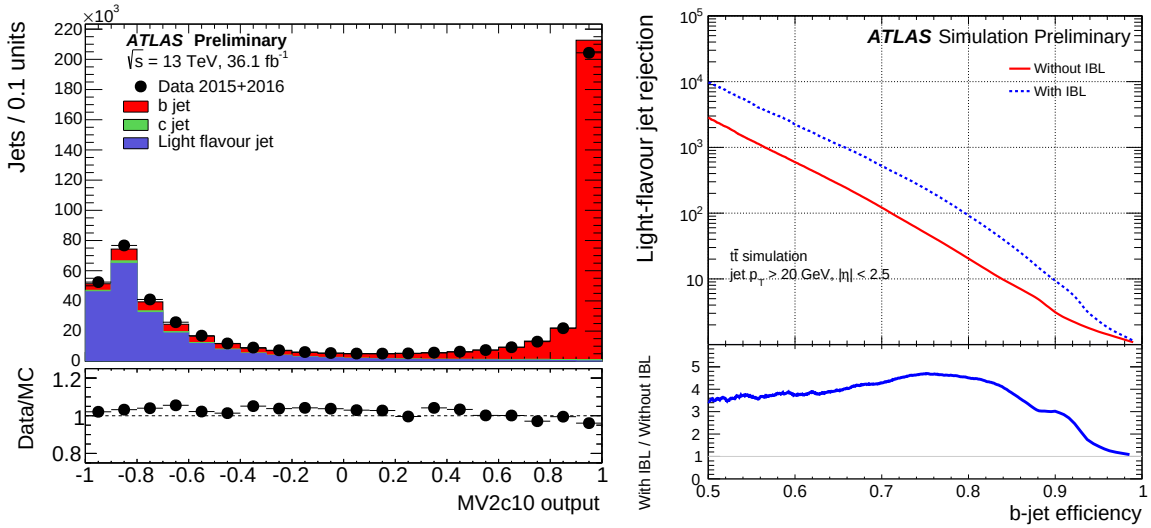
– **Secondary vertices:** The SV1 algorithm [130] looks for secondary vertices by taking tracks associated with the jet and that are significantly displaced from the primary vertex and forms vertex candidates from track pairs with the requirement $\chi^2_{vertex} < 4.5$. In a similar LLR to the IP3D algorithm, SV1 uses four variables to calculate a vertex weight: the invariant mass of all tracks in the vertex (assuming they are pions), the fraction of the total jet energy in these tracks, the number of two-track vertices, and the displacement in ΔR between the jet direction and a straight trajectory between primary and secondary vertices.

– **Decay chain:** The JetFitter algorithm [131] attempts to reconstruct the decay chain $b \to c \to s$ which has a distinct signature. Such a decay would result in a tertiary vertex as hadrons containing the charm have an appreciable lifetime just as for b-hadrons. A Kalman Filter is used to infer the path of the b-hadron by connecting the primary, secondary and tertiary vertices. Whilst this is effectively another example of a secondary vertex finder, it can provide increased sensitivity compared to SV1 as it exploits the unique topology of the $b \to c \to s$ decay chain.

The vertex-based tagging algorithms exhibit much lower mistag rates than impact parameter algorithms but their efficiency is limited by the secondary vertex finding efficiency. It is advantageous therefore to combine the two methods using an Multivariate Analysis (MVA) method which can take the correlations between variables into account. The MV2 algorithm [132] is a Boosted Decision Tree (BDT) which takes the variables used in the previously described taggers, along with the jet p_{T} and η as its inputs and outputs a single score which discriminates b-jets from light (u,d,s quark or gluon jets) and c-jets. This BDT is trained on approximately 5 million $t\bar{t}$ simulated events. The fraction of light-jets vs. c-jets that are designated

as background determines how well the discriminator will reject the two types of jets. The MV2c10 discriminant is trained using a ratio of light to c-jets of 90:10.

The distribution of the MV2c10 discriminant for simulated jets is shown in Figure 3.5a and shows clear discrimination between b-jets and other jets. The distribution for the BDT score in data is also shown and agrees well with the simulation. B-tagging algorithms benefit greatly from improved track and vertex resolution as one can see in Figure 3.5b, where the introduction of the IBL led to improved background rejection - hence the name 'Insertable B-Layer' [105].



(a) (b)

Figure 3.5: (a) The MV2c10 output shows good discrimination between jets originating from b-quarks and light jets [133]. (b) The improvement in b-Tagging performance after the introduction of the IBL [134]. For the same signal efficiency, the background rejection is far higher.

3.5 Muons

The relatively long lifetime - 2.2 μs - and highly penetrative properties of muons allow most of them to pass fully through the detector before decaying. A typical muon leaves an Inner Detector track, a small calorimeter deposit of energy and another track in the Muon Spectrometer (MS). This unique signature results in high identification efficiency and precise momentum measurements, which also makes muons excellent objects for triggering on events.

3.5.1 Muon Triggers

The ATLAS trigger has two stages, Level 1 (L1) and the High-Level Trigger (HLT), as described in Section 2.4. The L1 muon trigger system looks for coincident hits in the layers of the RPC or TGC systems. The geometrical coverage of the muon trigger system in the end-caps is $\approx 99\%$ but cables and equipment servicing the Inner Detector and calorimeters must pass through a small 'crack' in the MS at $\eta = 0$ which lowers the coverage in the barrel to $\approx 80\%$. The Region of Interest (RoI) of the muon candidate is used as a seed for the HLT algorithms where more advanced algorithms are executed.

The HLT muon reconstruction is divided into *fast* and *precision* stages, where the fast stage is designed to encourage early rejection of poorly reconstructed or fake muons and the precision stage uses algorithms as close to the full, offline reconstruction as possible [135]. The first, fast step is to refine the L1 muon candidate by including data from the MDT chambers in the RoI. A track fit is performed using the MDT drift times and positions and the p_T is estimated using look-up tables. This track is extrapolated to the Inner Detector where it is combined with tracks found there. The precision stage repeats these steps with finer granularity for any muon candidate that passes a set of quality criteria at the fast stage. Additional requirements can also be imposed during the precision stage for the muon trigger to accept the event. For instance, it is often required that the muon is *isolated* - that is to say the region around the muon candidate does not contain significant tracks or energy deposits. Muons can often be found in jets, where they are little more than by-products and only if the jet passes a jet trigger, should the event be stored. An isolation requirement is designed to reject these 'uninteresting' muons.

This analysis uses the logical OR of a low-p_{T} isolated trigger, for example, `mu_26_ivarmedium` and a higher p_{T} threshold unisolated trigger - `mu_50` [135]. The p_{T} threshold and isolation requirements for the lower p_{T} trigger varies with LHC run number to account for changes in running conditions. High p_{T} muons are more likely to deposit some energy in the calorimeters or leave low-energy tracks from bremsstrahlung radiation and therefore fail the isolation requirement. This combination of triggers makes use of the best features of them both, accepting lower p_{T} muons whilst maintaining efficiency at higher values of p_{T}.

The muon trigger efficiency can be measured using a *tag-and-probe* method looking at $Z \to \mu\mu$ events . These events are selected by triggering on one muon, and one can then study the efficiency of the trigger on the other muon. This procedure is performed both for data and simulated events and the small discrepancies can be used to derive scale factors that correct the simulated efficiency.

3.5.2 Muon Reconstruction

The full offline reconstruction of muons is very similar to that described above for the muon triggers [122]. However, with less strict restrictions on CPU usage and time taken, the offline reconstruction can maximise the precision of the momentum and track parameter measurements. Tracks in the ID and MS are reconstructed separately as described in Section 3.1 and the combination algorithms used depend on the information available in the ID, MS and calorimeters.

Four different muon types are defined based on which sub-detectors are used in the reconstruction:

- **Combined muon (CB):** A combined track is formed from a global refit of hits in the ID and MS. Most muons are reconstructed following an outside-in pattern recognition; extrapolating MS tracks back to the ID however a complementary, inside-out approach is also employed.

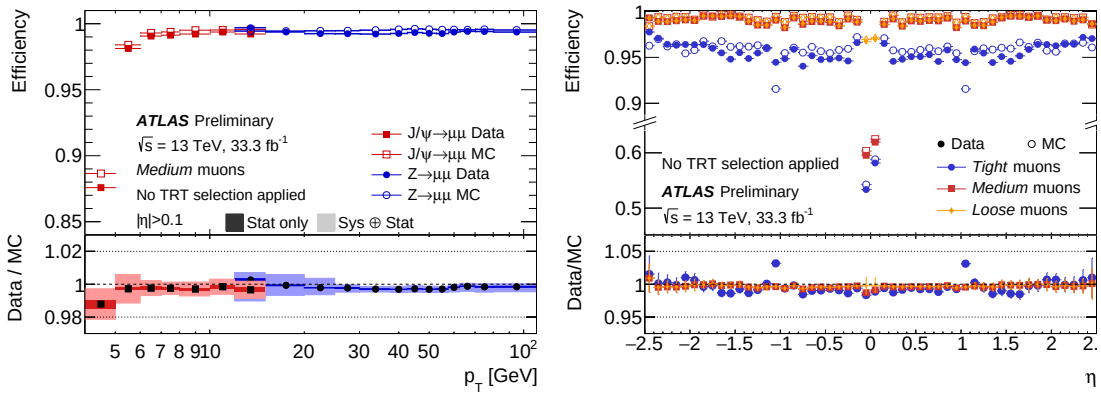
- **Segment-tagged muon (ST):** Muons that have a track in the ID with an associated track segment in the MDT or CSC chambers. ST muons are used when muons cross only one layer of the MS, either because of their low-p_{T} or because they pass through a region of reduced MS acceptance.

- **Calorimeter-tagged muon (CT):** A track in the ID is identified as a CT muon if it can be matched to an energy deposit in the EM calorimeter compatible with a minimally-ionising particle. This type of muon suffers from the lowest purity but recovers acceptance in regions where there are gaps in the MS such as the crack at $|\eta| < 0.1$.

- **Extrapolated muon (ME):** The reconstructed trajectory is based solely on track segments in the MS with a loose requirement that it originated from the interaction point, after extrapolating back to the centre of the detector and accounting for the expected energy losses in the calorimeters. ME muons are used to extend the acceptance for muon reconstruction out to $2.5 < |\eta| < 2.7$ which is not covered by the ID.

Overlaps between muon types are resolved before their use in physics analyses. Where different muon types share the same ID track preference is given to CB muons, then ST and finally CT muons. Overlaps with ME muons are resolved by selecting the track with the better fit quality and larger number of hits. For the analysis described in this thesis, only combined muons are used to ensure the best possible muon momentum resolution.

Muon reconstruction and identification efficiencies are measured in $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ events, using the same tag-and-probe method used to estimate the muon trigger efficiencies [122]. The muon reconstruction efficiencies as functions of p_{T} and η are shown in Figure 3.6a and 3.6b respectively. The reconstruction efficiency is relatively constant with p_{T} although the visible drops in efficiency for certain values of η reflect regions of reduced MS acceptance. Differences between simulation and data are used to derive correction factors of approximately 1-2% in magnitude.



(a) (b)

Figure 3.6: Muon reconstruction efficiency measured in $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ events as a function of p_{T} (a) and η (b) [136].

3.6 Other physics objects

3.6.1 Electrons and Photons

The typical electron signature is a track in the Inner Detector and a shower in the LAr calorimeter with no significant shower in the hadronic calorimeter. A photon will leave a similar shower but will not leave a track in the Inner Detector. In reality, it is not uncommon to observe so-called *converted* photons, where the photon interacts with the material in the Inner Detector and produces an electron-positron pair. It is therefore critical to distinguish 'real' electrons from converted photons.

The reconstruction algorithm that ATLAS employs [137] begins by using the topo-clusters described in Section 3.2. Since topo-clusters include both EM and hadronic calorimeters, and electrons/photons are expected to deposit almost all of their energy in only the EM calorimeter, a selection is made using the fraction of the cluster energy in the layers of the LAr calorimeter:

$$f_{EM} = \frac{E_{L1}+E_{L2}+E_{L3}+w\cdot(E_{E4}+E_{PS})}{E_{cluster}}, \quad w = \begin{cases} 1 & \text{for } 1.37 \leq |\eta| \leq 1.63 \\ 0 & \text{otherwise} \end{cases}, \tag{3.4}$$

where E_{Lx} corresponds to the energy in the x^{th} layer of the LAr calorimeter. In the transition region between the barrel and end-cap, $1.37 \leq |\eta| \leq 1.63$, a significant fraction of the total electron/photon shower energy is expected to be deposited in the presampler or a module of the tile-calorimeter known as the E4 *crack scintillator*. The energy deposits in these modules, E_{PS} and E_{E4} respectively, are therefore included in this region. A threshold of $f_{EM} > 0.5$ was chosen as it rejects 60% of pile-up clusters whilst maintaining good efficiency for true electron topo-clusters [137].

The next step is to match the topo-clusters with any tracks in the Inner Detector. Tracks that are within the RoI defined by the topo-cluster, which were reconstructed using the procedure described in Section 3.1, are refitted using a Gaussian Sum Filter (GSF), a non-linear generalisation of the Kalman filter which gives better track parameter estimation [138, 139]. To identify converted photons, a specialised photon-conversion vertex reconstruction algorithm is utilised, which looks for tracks emanating from a vertex consistent with a massless particle. The tracks from a converted photon can be reconstructed in just the TRT, the Silicon layers or both, and to complicate matters further it is possible for one of the tracks to be so soft ($p_{\mathrm{T}} < 0.5$ GeV) that it is not reconstructed. The reconstruction algorithm takes each of these cases into account.

To recover the energy from bremsstrahlung radiation, selected topo-clusters are used to seed unimaginatively-named *superclusters* where satellite clusters are combined with the seed cluster. Superclusters are initially designated as electron and/or photon candidates depending on the presence of suitable matching tracks. If the same topo-cluster has seeded an electron and photon supercluster, an ambiguity resolving algorithm will decide which hypothesis to reject after searching for additional tracks and conversion vertices and applying additional corrections and calibrations [137]. The finalised supercluster is then used to build an electron or photon object to be used in physics analyses.

An MVA approach is used to calibrate the energy measurements of photons and electrons, which uses the energies and pseudorapidities of contributing clusters as inputs to derive a correction factor that should be applied to the raw energy to obtain the 'true' energy.

There are many sources of uncertainty in the measurement of electron and photon energies. These include imperfect descriptions of the material in the Inner Detector, the response of different calorimeter layers and pile-up effects. These energy scale uncertainties are estimated as functions of p_{T} and η and verified by comparing $Z \to e^+e^-$ data events to simulated events with all calibrations applied. There is good agreement, within uncertainties, across the electron pair invariant mass spectrum, as shown in Figure 3.7.

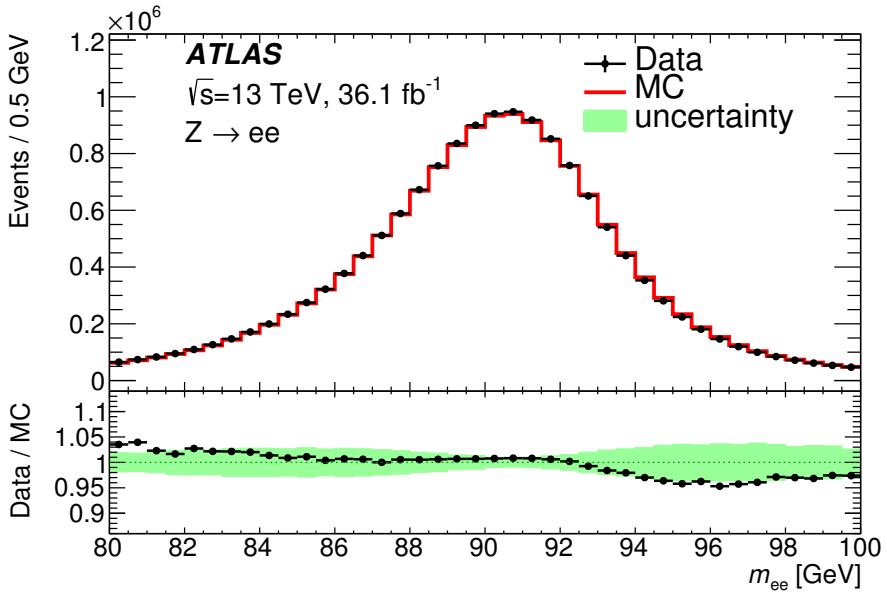


Figure 3.7: Electron pair invariant mass distribution from $Z \to e^+e^-$ decays in data compared to simulation after applying the full calibration. The lower panel shows the residuals for the data/MC ratios along with the total uncertainty (shaded green band). Discrepancies between data and simulation are well covered by the assigned uncertainties [140].

3.6.2 Taus

The tau lepton has a lifetime of 2.3×10^{-13} s, corresponding to a proper decay length of 87 μm. This means that most taus will decay long before reaching the first tracking layer, the IBL - 3.3 cm from the beamline - and only the tau decay products will be detected. The short lifetime of the tau is in part due to its large mass, 1.78 GeV, which allows it to decay into a multitude of hadronic final states. In fact, the branching ratio for hadronic decays is approximately twice as large as the leptonic decays $\tau^- \to e^-\overline{\nu}_e\nu_\tau$ and $\tau^- \to \mu^-\overline{\nu}_\mu\nu_\tau$.

The leptons from tau decays are virtually indistinguishable from prompt muon and electrons and so no special reconstruction algorithm is used to identify them. However, the jets from hadronically decaying taus can be reasonably well distinguished from other types of jet by the shape of the decay showers [141]. A final-state quark or gluon is much more likely to fragment before hadronising and forming a jet compared to the electroweak production of quarks from a tau decay. The higher probability of fragmentation in QCD jets also results in higher charged particle multiplicities and therefore more tracks associated with the jet. For a tau decay, charge conservation requires that the number of charged decay products is odd and hadronic final states are often categorised by the number of associated tracks or *prongs*. For instance, the decay $\tau^- \to \pi^-\nu_\tau$ would be a one-prong decay whereas $\tau^- \to \pi^-\pi^-\pi^+\nu_\tau$ is a three-prong decay.

Hadronically decaying tau candidates are reconstructed from jets, clustered with the anti-k_t algorithm with distance parameter, $R = 0.4$ and using the topo-clusters described in Section 3.2 as the inputs to the algorithm. These jets are required to have $p_{\mathrm{T}} > 10$ GeV and to be within the acceptance of the tracking detector, $|\eta| < 2.5$. Tracks within a cone of $\Delta R < 0.2$ are associated to the tau decay whereas tracks in the region $0.2 < \Delta R < 0.4$ are used to evaluate the isolation of tau candidate.

To identify tau candidates from other jets, BDTs are trained on shower-shape variables using $Z \to \tau^+\tau^-$ simulated events as signal and QCD multijet events as background [142]. One-prong and three-prong decays are trained separately and the resulting BDT scores gives good discrimination between tau jets and other jets, as can be seen in Figure 3.8. The optimal choice of BDT score threshold is dependent on the particular background and signal compositions of an analysis so three working points are chosen: 'loose', 'medium' and 'tight' each offering different levels of signal efficiency and background rejection.

The Tau Energy Scale (TES) is a correction factor to scale the reconstructed tau momentum back to the true tau momentum. It is calculated in bins of pseudorapidity and momentum and separately for one-prong and three-prong taus. The uncertainty on the TES is typically below 4%.

3.6.3 Missing Transverse Energy

Whilst the momenta of the individual partons is not known (see Section 2.1.1), the transverse momentum of protons in the LHC is negligible. One can therefore assume that the total transverse momentum in the event is vanishingly small. In most cases, the final-state momentum can be found by measuring particle tracks and

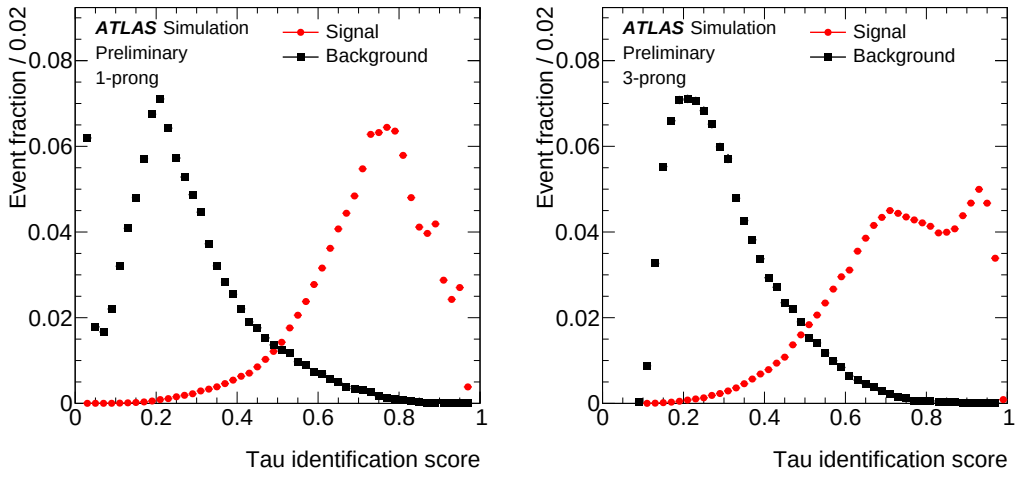


Figure 3.8: BDT scores for 1-prong (left) and 3-prong (right) tau identification. The signal taus are from simulated $Z \to \tau^+\tau^-$ events and the background are jets from QCD multijet simulated events [143].

calorimeter deposits although weakly-interacting particles, such as neutrinos, will carry away momentum without being detected. It is therefore useful to define the $E_T^{miss} = -\Sigma p_T$, a vector sum of the p_{T} of every reconstructed object in an event. The E_T^{miss} can be used as a proxy for the neutrino p_{T} in calculating kinematic quantities. In Dark Matter searches, the E_T^{miss} is often an important variable as, just like neutrinos, Dark Matter would not interact with the detector.

The E_T^{miss} is calculated in the x and y directions separately by combining the momenta of physics objects as well as tracks and calorimeter deposits that have not been associated with a particular physics objects, in the so-called *soft* term [144]:

$$E_{x(y)}^{miss} = -\Sigma p_{x(y)}^{jets} - \Sigma p_{x(y)}^{e} - \Sigma p_{x(y)}^{\gamma} - \Sigma p_{x(y)}^{\tau} - \Sigma p_{x(y)}^{\mu} + E_{x(y)}^{SoftTerm}. \tag{3.5}$$

There is some subtlety in the calculation of E_T^{miss} through the choice of which objects to include in the calculation. For instance, the ATLAS default is to use only jets clustered with the anti-k_t algorithm with $R = 0.4$ and $p_{\mathrm{T}} > 20$ GeV. The momenta of each physics object is calibrated using the techniques described previously before being used in the E_T^{miss} calculation.

3.6.4 Summary

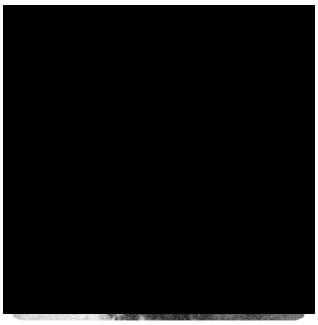
This section has described how signals in the detector can be combined to infer the passage of physics objects, measure their momenta and then calibrate this measurement. In particular, the reconstruction of muons and b-jets has been discussed in detail as it is these objects that form the experimental signature studied in this thesis.

Chapter 4

Signal and Background modelling

Fast is fine, but accuracy is everything

Xenophon



4.1 Simulation of events

In order to look for new physics one must first understand how its *signal* will appear in the detector and contrast this with the *background* Standard Model processes that will leave a similar signature. To do this, simulated events are generated but since Quantum Mechanics is frustratingly stochastic in nature it is necessary to generate a very large number of events. This is so that the distributions of important kinematic variables such as p_{T} and η are well modelled. The final goal of the simulation stage is to produce event samples in an identical format to data, such that both data and simulated samples can be treated identically in later stages of the analysis. This chapter shall begin by briefly describing the steps needed to produce a sample of simulated events. This will be followed by a more detailed overview of the signal

and background samples used in this thesis. Finally a validation that these samples give a good representation of real physical events will be given.

4.1.1 Parton Level

The first step in generating simulated events is to calculate matrix elements for all of the contributing processes using the technique described in Section 1.1.2; finding all of the Feynman diagrams contributing up to a particular order.

These matrix elements are functions of the momenta of the initial and final-state particles. There is a distribution of probabilities for the choices of momenta and for each event these distributions are sampled using a Monte Carlo (MC) technique [145]. Effectively, MC methods generate large quantities of random numbers and then select only those values which will re-create the desired distribution. The four-momenta of other particles can be determined from the conservation of momentum and energy.

In the case of the initial-state particles, the momenta of partons in the proton are described using Parton Distribution Functions (PDFs), important for the factorisation of the hard-scatter interaction in matrix element calculations described in Section 2.1.1. These PDFs are not calculable *a priori* and are instead found by fits to experimental data, particularly of Deep-Inelastic Scattering (DIS) processes. Figure 4.1 shows two PDF sets produced by the NNPDF collaboration [146] at two different energy scales, as functions of the longitudinal momentum fraction, x. The u_V & d_V peaks are for the valence up and down quarks respectively and all other bands are for *sea* quarks and gluons. At lower energies the uncertainties on the PDFs - the width of the bands - is larger, but for higher energies the PDFs are far more precisely known.

In order to compare a set of simulations with data, it is necessary to weight the simulated events such that they represent the correct number or fraction of a data event. These weights can be adjusted to tune distributions of kinematics to better match experimental data.

4.1.2 Parton shower and hadronisation

The result of the event generation stage are a number of final-state particles. These include quarks and gluons which are not directly detected, instead forming jets of particles. This process is often separated into two main stages. First the high-momenta partons will emit $q\bar{q}$ pairs in a so-called *parton shower*. Colour-confinement dictates that these quarks will then combine to form colourless hadrons. This hadro-

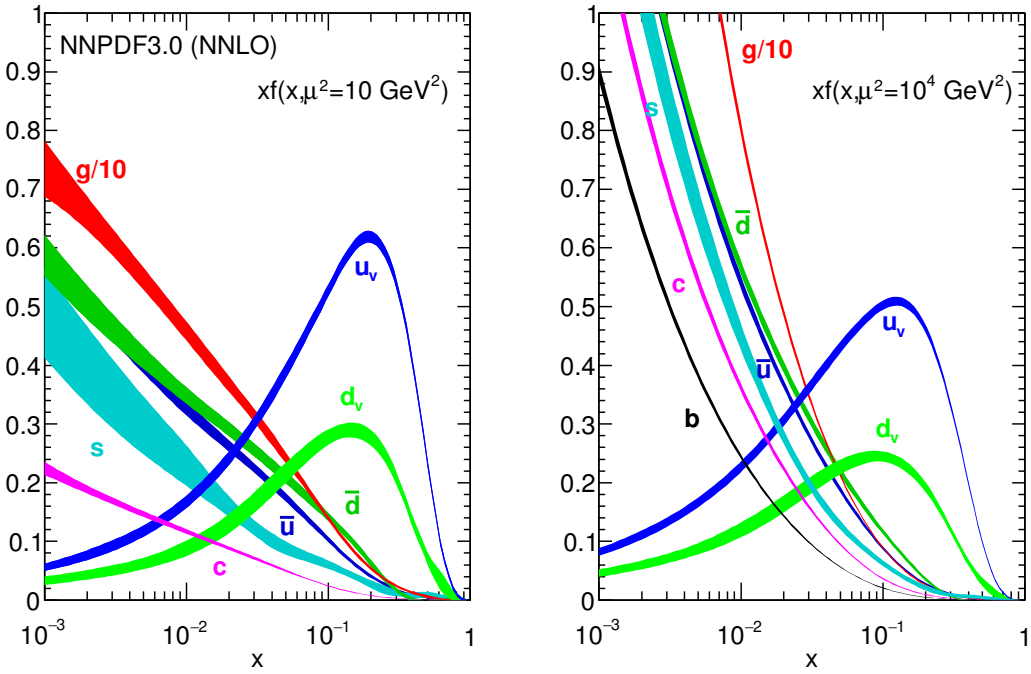


Figure 4.1: Parton Distribution Functions produced by the NNPDF collaboration [8].

nisation process is non-perturbative and therefore impervious to exact analytical solutions. Instead, there are two successful phenomenological models for parametrisation: the Lund String model and the cluster model which are used in the PYTHIA and HERWIG programs respectively [147, 148]. Each model attempts to describe the types and momenta of hadrons produced and has a number of non-physical parameters which can be tuned to match experimental data.

– **Lund String model:** In QCD, for large charge separation, colour field lines are compressed into tube-like regions modelled with $1+1$ dimensional strings. The strong force in such cases is approximately constant ($F(r) \approx \text{const} = \kappa \approx 1$ GeV$/fm$) and so the potential is linear $V(r) = \kappa r$. When there is enough potential energy in the colour field, a new hadron will be created. The fraction of energy given to this hadron is given by the fragmentation function which has parameters that can be fitted to experimental data. This fragmentation process continues until there is not enough energy to create new hadrons [149].

– **Cluster model:** Cluster hadronisation uses the *preconfinement* property of QCD [150] to find clusters of colourless groups of partons to form 'proto-hadrons' which can then decay into the final-state hadrons [151].

There is a subtlety in the use of parton showers, however. Consider an event that when reconstructed contains $N + 1$ jets. This can be achieved in one of two ways: firstly the collinear/soft radiation from showering an event containing $N+1$ partons, or an event with N partons with a hard, large-angle emission that leads to an extra jet. In this case, a jet can arise from either the matrix element or parton showering. To prevent double-counting, a matching scheme must be used to decide on an event-by-event basis which path should be taken. All simulated samples in this thesis use the CKWW-L matching scheme [152, 153] which defines a momentum cut-off k_T, above which jets should arise from the matrix element and below which, jets are simulated by the parton shower.

4.1.3 Detector Response

Once a final list of particles and their momenta is generated for each event, it is necessary to simulate how these particles will interact with the detector. A framework called GEANT4 [154, 155] is used to simulate the passage of particles through the detector, taking into account the exact geometry and material densities. Once GEANT4 has produced a set of *hits* where the particles have deposited energy in different regions of the detector, the next step is to simulate how these energies will be recorded by the electronic readout system. The *digitisation* step takes GEANT4 hits and converts them into voltages and currents whilst overlaying detector noise and pile-up events. The final step is to emulate the Read-out Drivers (RODs) - the hardware on the detector which converts the digitised hits into a data format that can be stored and used later in reconstructing events. As is shown in Figure 4.2, the objective is to create Raw Data Objects (RDOs) for simulated and data events so that they can be reconstructed in an identical fashion, as described in Chapter 3.

In the case of ATLAS specifically there are two different detector simulation schemes, one is called FULLSIM and uses GEANT4 as described above. However, approximately 80% of the full simulation time is spent simulating particles traversing the calorimeters and 75% is spent specifically simulating electromagnetic particles. This motivates the use of the FASTSIM scheme which replaces low-energy electromagnetic particles in the calorimeter with pre-simulated showers already loaded into memory [156]. This reduces computing time by a factor of three with little effect on the physics performance. The FASTSIM procedure is often used in events which do not require extremely accurate modelling of the calorimeters response and, in particular, electromagnetic shower shape.

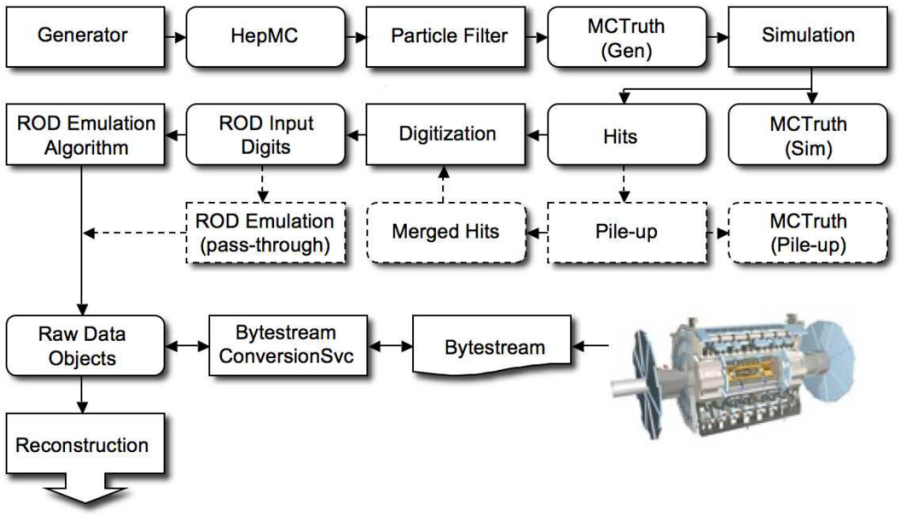


Figure 4.2: General layout of the ATLAS simulation framework. Starting from Monte Carlo generators, the goal is to produce Raw Data Objects which are reconstructed in an identical fashion to real data events. The important steps of generation, simulation of hits, and digitistation are discussed in detail here, the other steps correspond to the conversion of the data between different file formats and keeping track of meta-data generated during the simulation process [157].

4.2 Signal

Whilst this analysis is as model-independent as possible, a simulated sample of events must use some assumptions. In this case, signal samples were generated using a heavy Higgs model as a proxy. Samples were generated with a heavy Higgs width of 4.07 MeV under the Narrow width approximation (NWA) [158] for the two production mechanisms considered: b-associated production ($bb\Phi$) and gluon-fusion production (ggF) with Feynman diagrams as shown in Figures 1.8a and 1.8b respectively.

4.2.1 b-Associated production

The MadGraph5_aMC@NLO [159, 160] generator is used to generate $bb\Phi$ events in the 4-Flavour Scheme (4FS). This scheme assumes that the proton does not contain b-quarks and only diagrams such as in Figure 1.8a are allowed. Therefore only compatible Parton Distribution Functions (PDFs) can be used, and in this case the CT10nlo_nf4 PDF sets [161] were used to generate $bb\Phi$ events. For the parton showering, underlying event and hadronisation calculation Pythia 8.210 [147] was used with the A14 set of tuned parameters [162] and NNPDF2.3LO PDF sets [163].

The b-Tagging algorithms that attempt to identify jets from associated b-quarks use the properties of the b-hadron decay and so it is critical to model this process well. The EvtGen v1.2.0 program [164] was used to model the properties of the bottom and charm hadron decays. Details of the $bb\Phi$ samples generated are summarised in Table 4.1.

Samples were generated at nine mass points between 200 GeV and 1 TeV, in steps of 100 GeV. The $bb\Phi$ samples at 200, 400 and 1000 GeV were produced with much larger numbers of events and with both FullSim and FastSim reconstruction applied. This allowed the comparison of the two reconstruction schemes, and verified that the use of FastSim was appropriate for producing the remaining signal samples.

Table 4.1: The calculated cross-section times branching ratio ($\sigma \times BR$) and number of generated events for the nine aMC@NLO $bb\Phi$ samples.

Mass [GeV]	$\boldsymbol{\sigma \times BR}$ **(pb)**	$\boldsymbol{N_{evt}}$ **[k]**	**FullSim/FastSim**
200	8.446×10^{-02}	1000	Both
300	1.810×10^{-02}	50	FastSim
400	5.298×10^{-03}	1250	Both
500	1.889×10^{-03}	50	FastSim
600	8.389×10^{-04}	50	FastSim
700	3.950×10^{-04}	50	FastSim
800	1.915×10^{-04}	50	FastSim
900	1.050×10^{-04}	50	FastSim
1000	5.806×10^{-05}	1000	Both

4.2.2 Gluon fusion production

For the gluon fusion production mechanism, Powheg-Box v2 [165–168] was used - along with the CT10 PDF sets - in the matrix element calculation which was then interfaced with Pythia 8.210 with the AZNLO set of parameter tunes [169] and CTEQ6L1 PDF sets [170] to perform the showering and hadronisation. Although b-quarks are not required in the final state, as in the $bb\Phi$ case, it is still possible for Final State Radiation (FSR) to produce a bottom or charm hadron so, just as for the $bb\Phi$ samples, EvtGen v1.2.0 is used to model the decays of these hadrons. Samples were generated for the same nine mass points as the $bb\Phi$ samples and the details of these samples are summarised in Table 4.2.

Table 4.2: The calculated cross-section times branching ratio ($\sigma \times BR$) and number of generated events for the nine Powheg + Pythia8 ggF samples.

Mass [GeV]	$\boldsymbol{\sigma \times BR}$ **(pb)**	$\boldsymbol{N_{evt}}$ **[k]**	**FullSim/FastSim**
200	1.217×10^{-05}	50	FastSim
300	6.625×10^{-06}	50	FastSim
400	6.330×10^{-06}	50	FastSim
500	1.635×10^{-07}	50	FastSim
600	7.301×10^{-08}	50	FastSim
700	3.350×10^{-08}	50	FastSim
800	1.677×10^{-08}	50	FastSim
900	8.712×10^{-09}	50	FastSim
1000	4.733×10^{-09}	50	FastSim

4.3 Background

The major irreducible background processes are Drell-Yan production of $Z/\gamma^* + \text{jets}$ and the fully leptonic decay of $t\bar{t}$. There are also small contributions from Single Top and Diboson processes.

4.3.1 Drell-Yan $Z + \text{jets}$

Events containing Z/γ^* + jets (henceforth called $Z + \text{jets}$) were simulated with Powheg-Box v2 [167, 171], using the CT10 PDF sets and then interfaced to the Pythia 8.186 parton shower model [172]. This used the AZNLO tune and CT10 PDF set to model the non-perturbative effects. Bottom and charm hadron decays were modelled with EvtGen v 1.2.0 [164] and QED emissions from electroweak vertices and charged leptons were modelled using Photos++ 3.52 [173]. Since the dimuon mass distribution of $Z + \text{jets}$ is so strongly peaked at the Z boson pole mass, 91.2 GeV, only a small fraction of generated events will have a dimuon mass in the search region (0.2-1 TeV). For this reason, an additional set of $Z + \text{jets}$ samples are used with filters requiring a certain number of events in 'slices' of dimuon mass. These samples are summarised in Table 4.3.

Three mass-dependent scale factors were applied to correct for contributions from higher order terms. The first corrects the event yields to NNLO in the QCD coupling constant and is calculated using VRAP 0.9 [174] with the CT14NNLO PDF set [161]. Contributions from Electroweak (EW) and Photon-Induced (PI) ($\gamma\gamma \to \mu\mu$) processes at NLO were calculated using MCSANC 1.20 [175] and the MRST2004QED PDF set [176].

4.3.2 Top-antitop pair ($t\bar{t}$)

To generate $t\bar{t}$ events to NLO, the POWHEG-BOX v2 generator with the CT10 PDF sets was used. The parton shower, underlying event and hadronisation were simulated using PYTHIA 6.428 with the CTEQ6L1 PDF sets and the corresponding Perugia 2012 tune [177]. The top mass was set to 172.5 GeV and the sample was normalised to the predicted production cross-section as calculated with the TOP++ 2.0 program to NNLO in perturbative QCD which also includes soft-gluon resummation to next-to-leading-log order [178]. Details on the $t\bar{t}$ MC samples are also included in Table 4.3

Table 4.3: Summary of major background Monte Carlo samples including the cross-section times branching ratio ($\sigma \times BR$) and number of Monte Carlo events.

Process	**Mass [GeV]**	$\boldsymbol{\sigma \times BR}$ **(pb)**	$\boldsymbol{N_{evt}}$ **[k]**
$Z/\gamma^* \to \mu\mu$	> 60	$1.901 \times 10^{+03}$	2000
$Z/\gamma^* \to \mu\mu$	120-180	$1.748 \times 10^{+01}$	500
$Z/\gamma^* \to \mu\mu$	180-250	$2.921 \times 10^{+00}$	250
$Z/\gamma^* \to \mu\mu$	250-400	$1.082 \times 10^{+00}$	150
$Z/\gamma^* \to \mu\mu$	400-600	1.955×10^{-01}	100
$Z/\gamma^* \to \mu\mu$	600-800	3.740×10^{-02}	150
$Z/\gamma^* \to \mu\mu$	800-1000	1.061×10^{-02}	50
$Z/\gamma^* \to \mu\mu$	1000-1250	4.258×10^{-03}	50
$Z/\gamma^* \to \mu\mu$	1250-1500	1.422×10^{-03}	50
$Z/\gamma^* \to \mu\mu$	1500-1750	5.452×10^{-04}	50
$Z/\gamma^* \to \mu\mu$	1750-2000	2.299×10^{-04}	50
$Z/\gamma^* \to \mu\mu$	2000-2250	1.039×10^{-04}	50
$Z/\gamma^* \to \mu\mu$	2250-2500	4.94×10^{-05}	50
$Z/\gamma^* \to \mu\mu$	2500-2750	2.45×10^{-05}	50
$Z/\gamma^* \to \mu\mu$	2750-3000	1.25×10^{-05}	50
$Z/\gamma^* \to \mu\mu$	3000-3500	1.00×10^{-05}	50
$Z/\gamma^* \to \mu\mu$	3500-4000	2.90×10^{-06}	50
$Z/\gamma^* \to \mu\mu$	4000-4500	9.00×10^{-07}	50
$Z/\gamma^* \to \mu\mu$	4500-5000	3.00×10^{-07}	50
$Z/\gamma^* \to \mu\mu$	> 5000	1.00×10^{-07}	50
$t\bar{t}$	-	$3.783 \times 10^{+02}$	10000

4.3.3 Other backgrounds

Other backgrounds that could result in a similar signature to signal events are Diboson processes and single top-quark production. Whilst the contributions of these processes are fall smaller than $Z+$jets and $t\bar{t}$, they are included to give a more complete background description. The 'single-top' processes are modelled using the same procedure as the $t\bar{t}$ sample described in Section 4.3.2 and normalised using an approximate calculation at NLO in QCD for the s and t channels [179, 180] and a NLO+NNLL predictions were used for the Wt channel [180].

Diboson processes were modelled using the SHERPA 2.1.1 generator [181] and the contribution from one (ZZ) or zero (WZ, WW) additional partons at NLO and up to three additional partons at LO using the Comix [182] and OpenLoops [183] matrix element generators. These were merged with the SHERPA parton shower using the ME+PS@NLO prescription [184]. The CT10 PDF set was used in conjunction with a dedicated parton shower tune developed by the SHERPA authors. The diboson sample event yields were simply normalised to the generated cross-sections calculated at NLO.

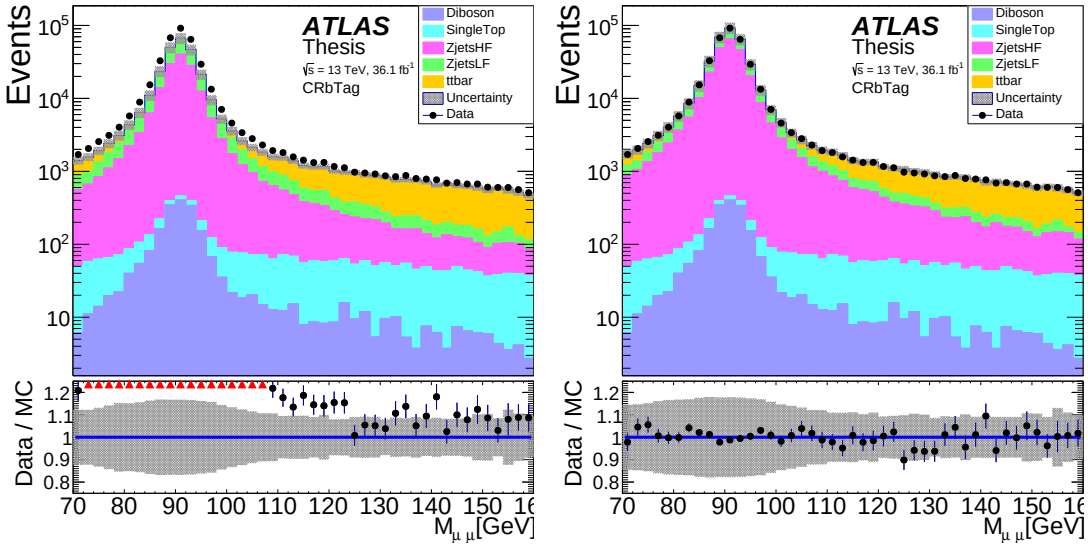
4.4 Validation

In order to effectively search for New Physics it is necessary to understand the expected background events. Distributions of kinematic variables can be compared between simulation and data to check that the Monte Carlo models describe these variables well. Events are chosen with similar selection criteria to the search regions which will be described in more detail in Section 5.2. These so-called Control Regions (CRs) cover a lower mass range ($m_{\mu\mu} < 160$ GeV) however, so looking at data in these regions should not bias the analysis later on.

4.4.1 Heavy flavour modelling in $Z+$jets

It is important to study the fraction of background events that contain b-quarks since this will be a key requirement in the search for the b-associated production mechanism. Figure 4.3a shows a comparison of the dimuon invariant mass spectrum between data and simulated events when requiring one b-tagged jet. There is a discrepancy about the Z resonance peak which indicates that the number of expected $Z+$jets events is too low. Figure 4.4 shows the number of b-jets produced by different Monte Carlo generators. One can see that Powheg differs in its prediction of the number of b-jets from the other generators. To correct for this discrepancy,

the $Z+$jets background sample is sub-divided. Events containing a c-quark or b-quark are designated as Z + Heavy Flavour (Z+HF) and all others are called Z + Light Flavour (Z+LF). Then by scaling up the Z+HF component, a good agreement between simulation and data can be achieved. In Figure 4.3b, a naïve Normalisation Factor (NF) of 1.5 is applied to the Z+HF background to demonstrate how this correction improves the modelling of the $m_{\mu\mu}$ distribution. In the full analysis this NF is determined by a fit of the expected background distribution to measured data.



(a) (b)

Figure 4.3: Discrepancy of $m_{\mu\mu}$ modelling around the Z resonance in events with at least one b-tag before (a) and after (b) applying a naïve Normalisation Factor of 1.5 to the Z+HF component of the background. The full selection used for this control region, CRbTag is defined later in Section 5.2. Red triangles represent data points outside of the frame.

4.4.2 Correlation of jet multiplicity and $p_T^{\mu\mu}$

Figures 4.5a and 4.5b show the leading and subleading muon p_{T}, where the MC background underpredicts the number of events with high p_{T} muons. The POWHEG $Z+$jets sample considers only one extra parton at NLO and so for jet multiplicities of two or greater, jets are only simulated by the parton shower. Just as for the b-jet modelling checks shown in Figure 4.4, the POWHEG sample was compared with other generators and discrepancies in the number of expected jets was observed. Since Z^0 bosons produced in association with jets will recoil, an underprediction of the jet multiplicity leads to an underprediction of the p_{T} of the Z^0 boson.

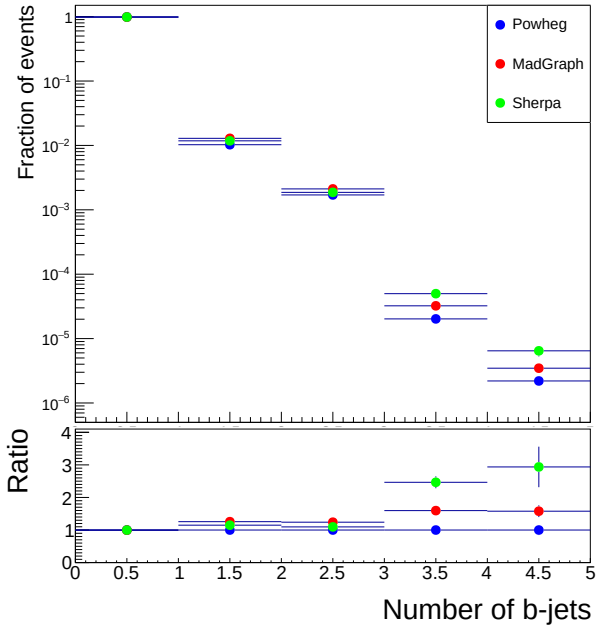


Figure 4.4: Differences between Monte Carlo event generators description of the number of b-jets in $Z+$jets events. The lower panel shows the ratio with respect to the Powheg MC samples. Higher b-jet multiplicities are predicted by the Sherpa and MadGraph MC generators.

To verify that the mismodelling of the lepton p_{T} is due to the insufficiencies in the modelling of the Z p_{T}, a check was performed where the $Z+$jet events were reweighted according to the p_{T} of the dimuon system, $p_T^{\mu\mu}$. To achieve this, the data/MC ratio was fitted as a function of the momentum of the dimuon system with a second-order polynomial. The fitted function was then used to obtain a correction factor for a given value of $p_T^{\mu\mu}$. This was then applied to each event and the effect on the modelling of other kinematic variables observed. The 'corrected' lead and sublead muon p_{T} spectra are shown in Figure 4.5, and one can see that this reweighting has significantly improved the modelling of the lepton p_{T}. This is unsurprising as the p_{T} of the individual muons is highly correlated with the p_{T} of the dimuon system.

The effect of this reweighting procedure on the $m_{\mu\mu}$ is small, however, as one can see from Figure 4.6 which shows that the change in shape of the dimuon mass distribution is smaller than the considered systematic uncertainties, described in detail in Chapter 7. One can therefore infer that the $p_T^{\mu\mu}$ and $m_{\mu\mu}$ are not strongly correlated and the insufficiencies in the Powheg modelling have little effect on the dimuon mass distribution which is used to search for a signal.

This small study provides a good explanation of the cause for discrepancies in the $Z+$jets modelling and validates that the systematic uncertainties considered in the analysis are adequate to encompass these effects. This somewhat informal reweighting procedure was used to visualise the impact of the $p_T^{\mu\mu}$ mismodelling but is not used later in the final analysis as this effect is dealt with by the considered systematic uncertainties described later in Chapter 7.

4.4.3 Modelling of $t\bar{t}$

The $t\bar{t}$ background provides the largest contribution to the predicted event yield for events containing b-Tagged jets, as can be seen from the tables in Appendix C. It is therefore critical that this background process is well modelled by the MC simulated samples. Figure 4.7 shows comparisons between the expected kinematic distributions with those measured in data for the $t\bar{t}$ control region, CRttbar, defined later in Section 5.2. In regions dominated by $t\bar{t}$ there is very good agreement and therefore the modelling is very good. There is a small discrepancy at $m_{\mu\mu} \approx 91$ GeV where the Z+HF background becomes non-negligible. Also, one can see that for muons with very high p_{T} there are only a small number of $t\bar{t}$ events. The statistical uncertainty of the MC prediction will therefore be higher in these regions.

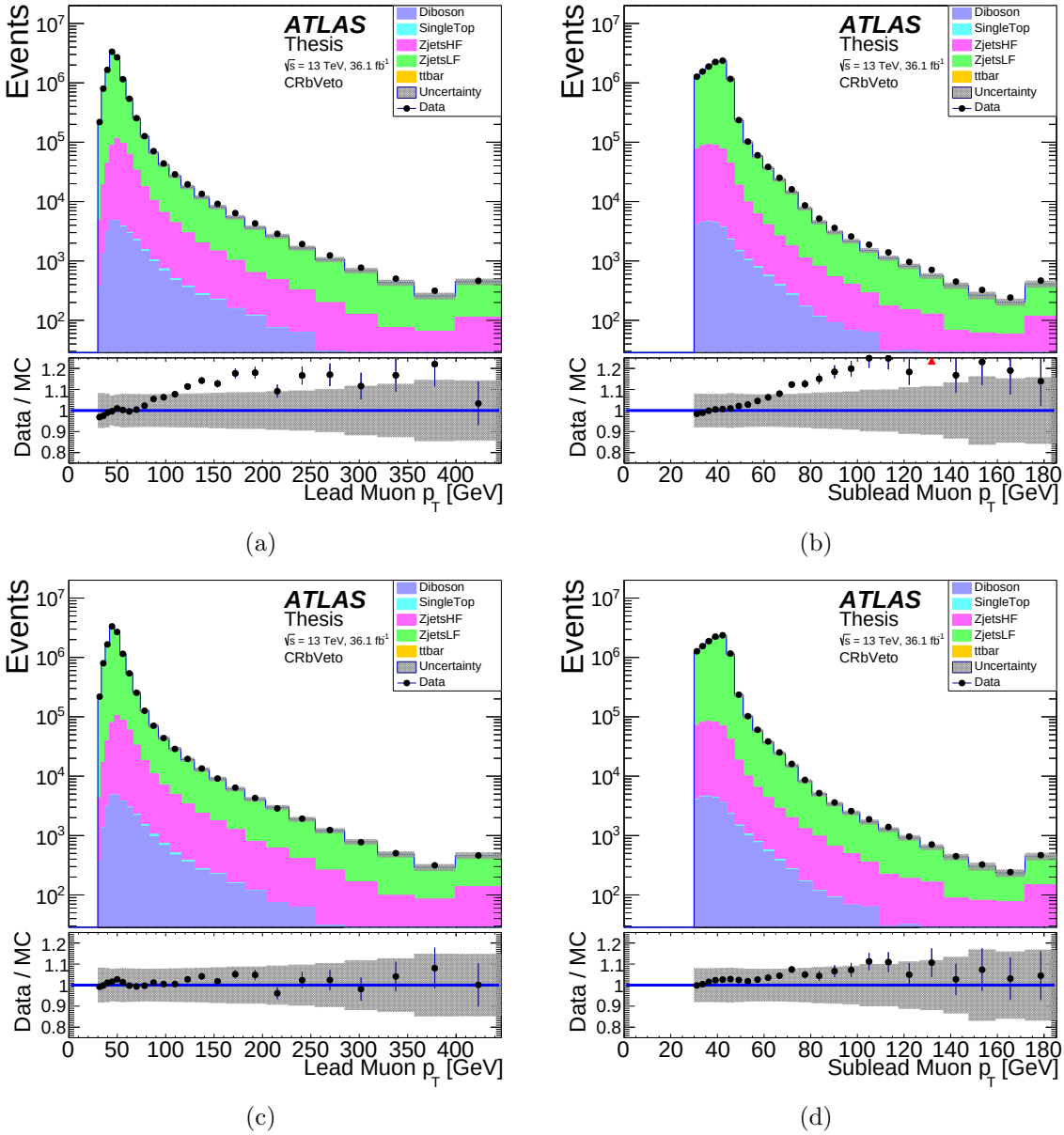


Figure 4.5: The lead (left) and sublead (right) muon p_{T} distributions in the CRbVeto control region (see Section 5.2) before (top) and after (bottom) applying a simple reweighting in $p_T^{\mu\mu}$. The points show observed data and the shaded bands represent the total pre-marginalised systematic uncertainties. Red triangles represent data points outside of the frame.

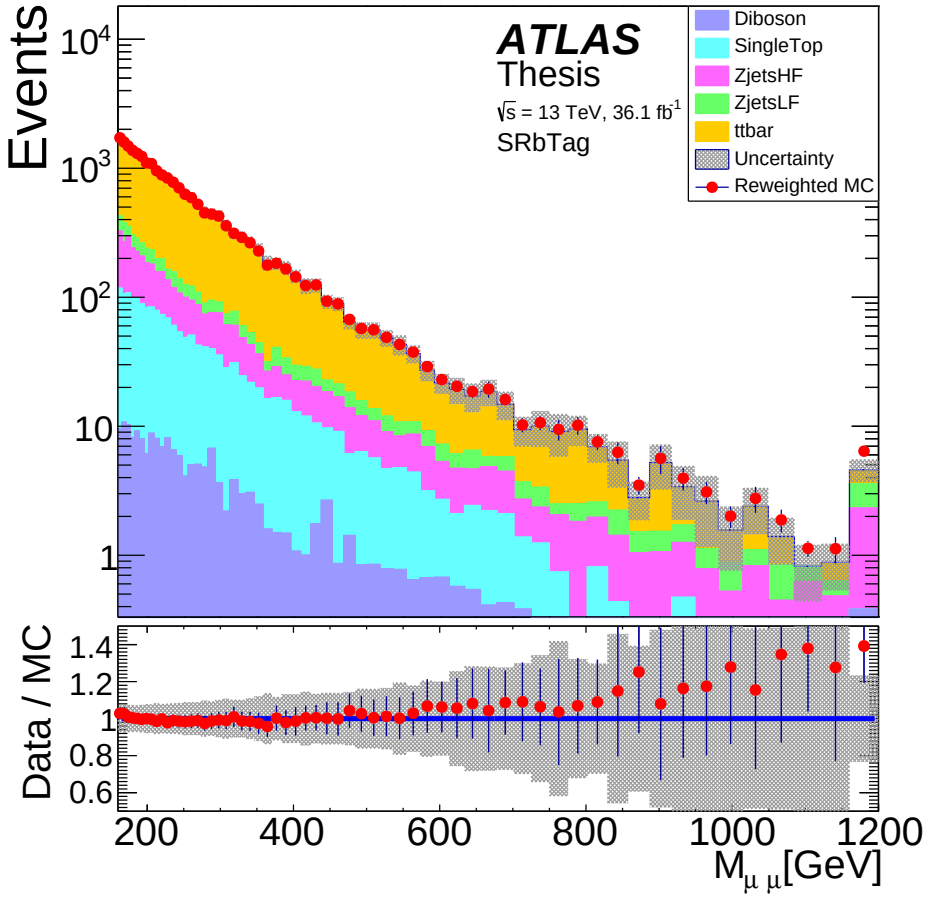
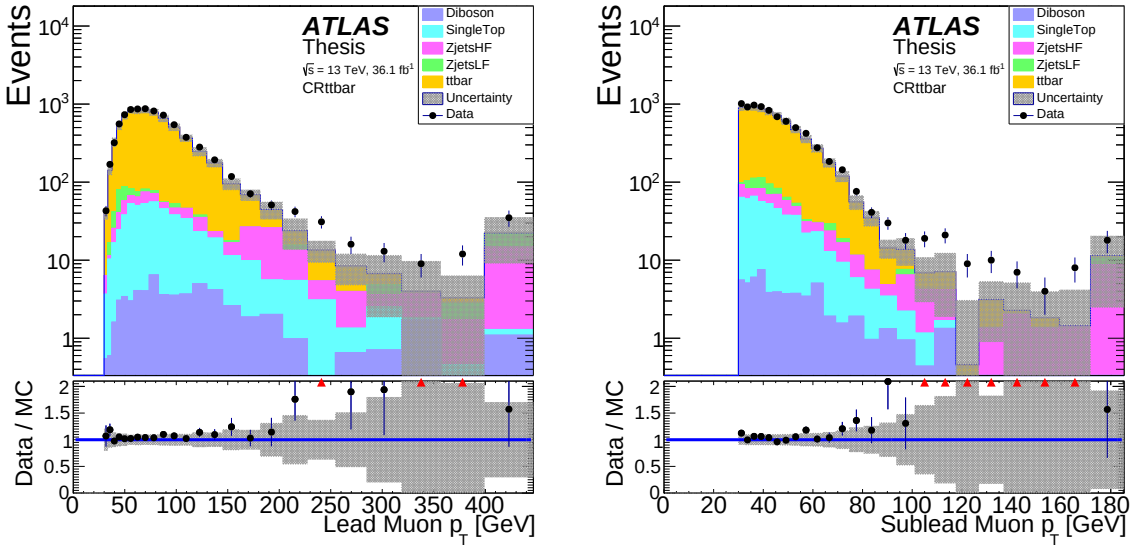
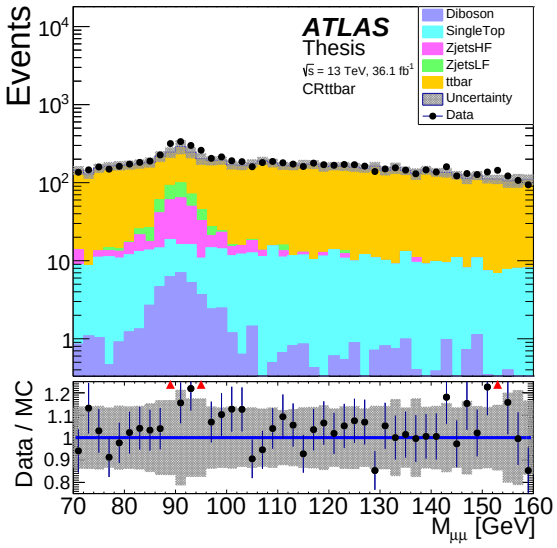


Figure 4.6: The effect of the p_{T} reweighting procedure on the predicted $m_{\mu\mu}$ shape of the background in the SRbTag signal region (see Section 5.2) is smaller than the pre-marginalised systematic uncertainties applied in this analysis.

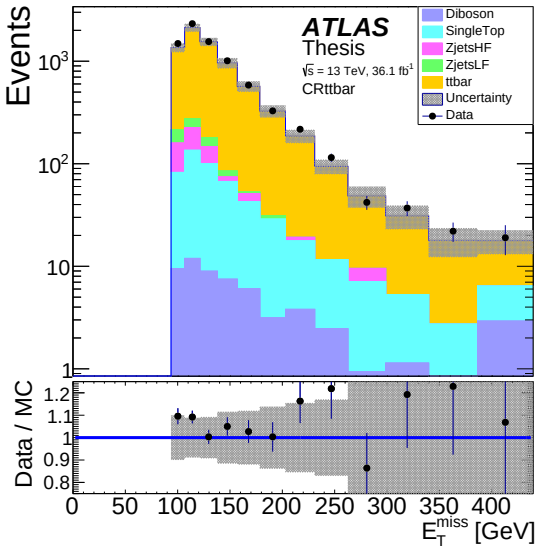


(a)

(b)



(c)



(d)

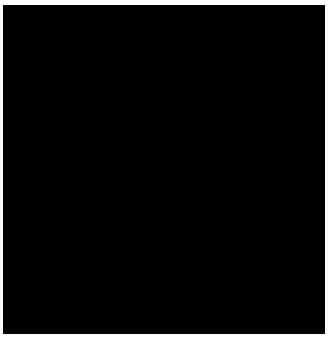
Figure 4.7: The lead (a) and sublead (b) muon p_{T}, dimuon invariant mass (c) and E_T^{miss} (d) distributions in the $t\bar{t}$ control region, CRttbar (see Section 5.2). Red triangles represent data points outside of the frame.

Chapter 5

Analysis Selections

It is our choices that show what we truly are, far more than our abilities.

Albus Dumbledore



The selection is separated into two stages; firstly the object *pre-selection*, where requirements are placed on which objects are considered in the analysis, and then the event selection where events are selected if they contain a chosen number and topology of objects. This chapter shall begin by defining the pre-selection, designed to ensure that only well-reconstructed objects - in particular muons and b-jets- are used in this analysis. Then the event selections will be described. These define regions that are useful either because they have the potential to contain signal events, so-called Signal Regions (SRs), or those which can be used to find normalisations of the SM background processes without signal contamination, known as Control Regions (CRs) .

5.1 Object Pre-selection

5.1.1 Muons

Reconstructed muon candidates are required to have a p_{T} greater than 30 GeV and pseudorapidity $|\eta| < 2.5$. The latter cut requires that the muon trajectory lies within the acceptance of the Inner Detector whilst the former ensures muons are well above the trigger p_{T} threshold, such that the trigger efficiency is constant and well modelled. Muons are required to be consistent with the hypothesis that they originate from the primary vertex by applying selections on the transverse (d_0) and longitudinal (z_0) impact parameters defined with respect to the primary vertex position: $|d_0/\sigma_{d_0}| < 3$ and $|z \sin\theta| < 0.5$ mm where σ_{d_0} is the uncertainty on the transverse impact parameter. Nearby tracks could negatively affect the reconstruction, so an *isolation* requirement is applied on the total momenta of tracks in a variable cone size of $\Delta R = \min(0.3,\ 10\ \mathrm{GeV}/p_{\mathrm{T}}^{\mu})$, centred on the muon track. This isolation requirement is tuned to yield a 99% efficiency over the full muon p_{T} spectrum [122].

Muons produced by a high-mass object are expected to have very high p_{T} and in these cases even the high magnetic fields in ATLAS will only slightly deflect the muon's path. This makes the measurement of the momentum challenging. To combat this, extra requirements are placed on the muon to minimise the momentum resolution for high-p_{T} muons. Muons are required to leave three hits in each of the three MS precision layers. A veto is placed on muons passing through poorly-aligned regions of the MS or those passing through the overlap region between the MS barrel and end-cap ($1.01 < |\eta| < 1.10$). These additional selections are approximately 80% efficient for high-p_{T} muons.

5.1.2 Jets

This analysis uses jets reconstructed from noise-suppressed topological clusters using the anti-k_t algorithm as described in Section 3.3 with radius parameter, $R = 0.4$. All jets are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 4.5$. Jets with p_{T} below 60 GeV and $|\eta| < 2.4$ must further satisfy a requirement on the Jet-Vertex Tagger (JVT) - a likelihood discriminant that uses track and vertex information in order to suppress jets originating from pile-up activity [185]. To prevent poorly measured jet energies from impacting other related measurements - in particular E_T^{miss} - events are rejected if they contain any jets that fail a "clean jet" requirement [186] which vetoes jets that are likely to have been reconstructed erroneously. An example would be a jet reconstructed from a sporadic noise burst in a calorimeter.

5.1.3 Overlap Removal

An overlap removal procedure is applied to prevent a single particle being reconstructed as two different physics objects. The first step is to remove overlaps between muons and jets. Jets not tagged as b-jets, that are reconstructed within $\Delta R = 0.2$ of a muon are removed if they have fewer than three associated tracks or if the muon energy constitutes more than 70% of the jet energy. Muons reconstructed within a variable cone size of $\Delta R = \min(0.4,\ 0.04 + 10\ \text{GeV}/p_{\text{T}}^{\mu})$ of any surviving jet are removed. Jets are also discarded if they lie within $\Delta R = 0.2$ of an electron candidate. In this case, electrons are reconstructed as described in Section 3.6.1 and are required to have transverse energy $E_T > 10$ GeV and pseudorapidity $|\eta| < 2.47$. The likelihood-based selection described in Ref. [140] is used at the "loose" operating point.

5.2 Event Selection

5.2.1 b-Tagging Requirements

The multivariate technique described in Section 3.4 is used to identify jets originating from b-hadrons. The `MV2c10` discriminant is the output of a Boosted Decision Tree (BDT) which has been trained to identify b-jets against a background sample containing an admixture of 10% charm jets and 90% jets originating from light quarks [132]. There are four supported operating points for the choice of `MV2c10` threshold, where the necessary corrections, calibrations and uncertainties have been estimated. Each ATLAS analysis must decide which operating point is best for their particular circumstances. In this search the "loosest" requirement, for which 85% of b-jets pass the selection, was chosen as this was found to provide the optimal signal sensitivity. An event is defined as 'b-tagged' if it contains at least one b-tagged jet that passes this threshold.

5.2.2 Signal and Control Regions

The signal regions are defined with the selections given in Table 5.1. The search region is for dimuon invariant masses above 160 GeV, this is the upper mass bound considered in the SM $H \to \mu\mu$ analysis [93]. The b-Tag region SRbTag requires at least one b-Tagged jet whereas the b-Veto region contains events with no b-tagged jets. In b-associated production, the b-quarks can often pass outside the geometric acceptance of the Inner Detector and therefore the complementary region SRbVeto has acceptance for the signal.

The control regions are defined with the selections given in Table 5.2. All control regions are below the 160 GeV invariant mass threshold. A complementary region to the SRbVeto signal region is defined with a similar veto on b-Tagged jets. A region designed to have a large contribution of $t\bar{t}$ events, CRttbar, is defined with a requirement on the E_T^{miss} as well as a minimum of one b-tagged jet. The CRbTag region has the same b-Tag requirement and an orthogonal cut on the E_T^{miss}. The expected event yields of the background and signal processes are given in the tables in Appendix C. The event yield for the SM Higgs decay process was also estimated from simulated samples generated for the $H \to \mu\mu$ analysis [93] and found to be negligible.

Table 5.1: Definitions of Signal Regions where $m_{\mu\mu}$ is the reconstructed dimuon invariant mass and $nBJets$ is the number of b-jets as defined in Section 5.2.1.

Region	**Definition**
SR bTag	$m_{\mu\mu} > 160$ GeV & $nBJets > 0$
SR bVeto	$m_{\mu\mu} > 160$ GeV & $nBJets = 0$

Table 5.2: Definitions of Control Regions where $m_{\mu\mu}$ is the reconstructed dimuon invariant mass, $nBJets$ is the number of b-jets as defined in Section 5.2.1 and E_{miss}^T is the missing transverse energy.

Region	**Definition**
CR bVeto	$100 < m_{\mu\mu} < 160$ GeV & $nBJets = 0$
CR bTag	$100 < m_{\mu\mu} < 160$ GeV & $nBJets > 0$ & $E_T^{miss} < 100$ GeV
CR ttbar	$100 < m_{\mu\mu} < 160$ GeV & $nBJets > 0$ & $E_T^{miss} > 100$ GeV

5.3 MVA Selection

Whilst the aim of the analysis is to perform a model-independent search, such that the results are useful for testing many different theories, it is possible to gain greater sensitivity to particular models by further optimizing the selection. Simply selecting on variables which allow discrimination between signal and background can improve sensitivity somewhat but often even more can be achieved through the use of a Multivariate Analysis (MVA), which can take correlations between variables

into account. This section describes studies into the potential improvements to a similar, future analysis that can be achieved through such a technique. The MVA selection used in this section is not used elsewhere in the analysis as these studies were performed after the results in Section 8 were finalised.

5.3.1 Introduction to Boosted Decision Trees (BDTs)

Boosted Decision Trees (BDTs) are a class of powerful multivariate techniques that are widely used in high-energy physics, particularly for classification applications [142, 187]. A decision tree is constructed from *nodes* and *leaves* such as in Figure 5.1. At each node the data is split based upon some feature - in the case of particle physics this is often a kinematic variable such as p_{T}. Each split is followed either by another node, or a leaf which gives the classification of the corresponding data. The decision at each node can be optimised to maximise the *information gain*, or more quantitatively by minimising the *entropy* given by [188]:

$$H(X) = -\sum_{i=1}^{n} p_i \log_2(p_i), \tag{5.1}$$

where n is the number of children for each node and p_i is the probability of each outcome. A single decision tree is often called a *weak* learner as they have high misclassification rates, but many such decision trees can be combined in a process called *boosting* to provide a classifier with much higher accuracy [189]. In these cases trees are created iteratively, and the output of each tree $h(x)$ is given a weight w related to its accuracy such that the ensemble output is:

$$\hat{y}(x) = \sum_{t} w_t h_t(x). \tag{5.2}$$

The goal of machine learning with BDTs is then to minimise the objective function:

$$O(x) = \sum_{i} l(\hat{y}_i, y_i) + \sum_{j} \Omega(f_j), \tag{5.3}$$

where $l(\hat{y}_i, y_i)$ is the loss function - the difference between the prediction and truth of the ith sample - and $\Omega(f_j)$ is the regularisation function which penalises the complexity of adding the jth tree. There are a myriad of choices for the loss and regularisation functions as well as methods to perform the minimisation, each with their own advantages and disadvantages [189–191].

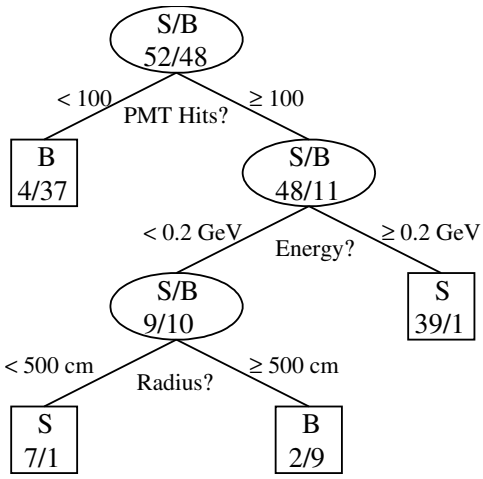


Figure 5.1: Schematic of an example decision tree used by the `MiniBooNE` experiment for particle identification (PID). The nodes are shown as ovals and the leaves as rectangles. The discriminating variables used are specific to the experiment: PMT hits, Energy and Radius. Each leaf is labelled with whether it should be classified as signal (S) or background (B). The number of simulated signal and background events at each node are shown in the format S/B [192].

5.3.2 BDT Training

A BDT was created in the `Scikit-learn` framework [193] using an `AdaBoost` classifier [189] with the hyperparameters and kinematic variables given in Table 5.3. This BDT was trained to identify $t\bar{t}$ events from the $Z+$jets and $bb\Phi$ signal samples in the b-Tag signal region described above in Section 5.2.2. The `Max_depth` and `Min_samples_leaf` parameters control the creation of each decision tree; the maximum depth of the tree, and the minimum number of samples required on each leaf. The `n_estimators` and `learning_rate` parameters control the boosting process through the number of weak learners to be trained and the initial weight of each classifier.

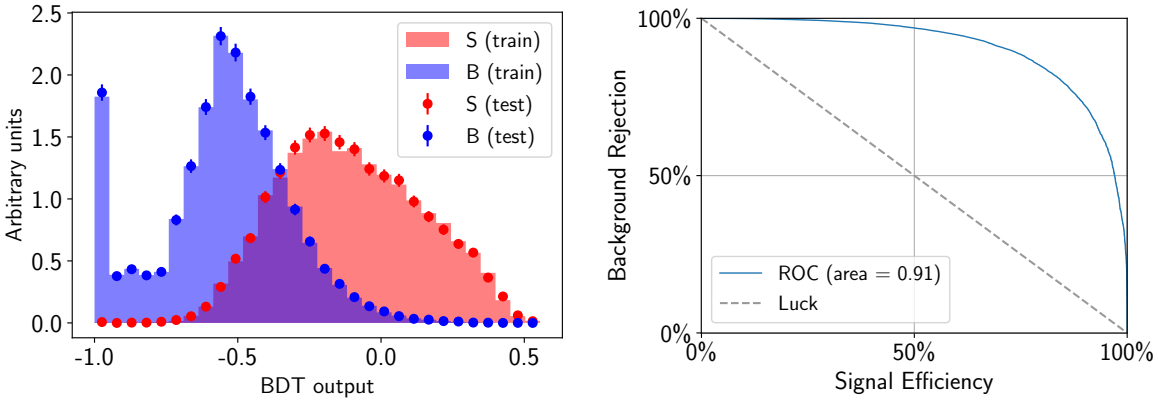
A major concern in machine learning is *overtraining* where the algorithm learns to distinguish samples by small statistical fluctuations in the dataset, resulting in spuriously high discriminating power. This effect can be measured by subdividing the datasets into *training* and *testing* samples, where the former is used for training the BDT and then compared to the BDT results for the latter. If there are significant differences between the two, then the BDT is overtrained. Figure 5.2a shows the BDT score distributions for the "signal" - which in this case is the $t\bar{t}$ background - and "background" which includes the $Z+$jets and $bb\Phi$ samples. One can see that there is good discrimination, and the agreement between the train and test samples indicates that overtraining is not a problem.

Table 5.3: Details of the BDT including the hyperparameters and kinematic variables.

Hyper Parameter	**Value**
`Max_depth`	7
`Min_samples_leaf`	100
`n_estimators`	1500
`learning_rate`	0.05
Kinematic Variables	1) Lead Muon p_{T} 2) Sub-lead Muon p_{T} 3) ϕ separation of dimuon pair, $\Delta\phi_{\mu\mu}$ 4) η separation of dimuon pair, $\Delta\eta_{\mu\mu}$ 5) Minimum angular distance between any muon and any b-jet, $\Delta R_{\mathrm{min}}(\mathrm{muon}, \mathrm{b\text{-}jet})$ 6) Missing transverse energy, E_T^{miss} 7) No. of b-jets passing 85% b-Tag operating point 8) No. of b-jets passing 77% b-Tag operating point 9) No. of b-jets passing 70% b-Tag operating point 10) No. of b-jets passing 60% b-Tag operating point

For a given BDT score, the fraction of signal greater - and background lower - than this threshold are called the *signal efficiency* and *background rejection* respectively. Figure 5.2b shows these fractions as a function of the BDT score, known as a Receiver Operating Characteristics (ROC) curve. A good BDT discriminant should maximise the area under this ROC curve, maximising the signal efficiency for a given background rejection and vice versa. For this BDT the area under the ROC curve is approximately 0.9 which quantifies the good discrimination one can see in Figure 5.2a.

The dimuon invariant mass, $m_{\mu\mu}$, was intentionally not included as a feature on which to train the BDT, as this variable is later used to search for the resonance peaks of a signal. It is possible, however, that the BDT has inadvertently been trained to use the $m_{\mu\mu}$ through its correlation with other variables. To study this, the correlation between the BDT score and the $m_{\mu\mu}$ was quantified. Figure 5.3 shows heat-maps of the BDT score and $m_{\mu\mu}$ for the $Z+$jets and $t\bar{t}$ MC samples as well as the combination of all $bb\Phi$ signal samples. One can also see the Pearson and Spearman's rank correlation coefficients, ρ_P and ρ_S, which quantify the degree of correlation [194, 195]. There is no correlation between the BDT score and $m_{\mu\mu}$ for the $Z+$jets and $t\bar{t}$ samples and only minor correlation ($\rho \approx 0.2$) in the signal samples. Thus, the discriminating power of the BDT does not come about because it has 'cheated' and learned to recognise events with large dimuon invariant masses.



(a) (b)

Figure 5.2: (a) BDT score distribution for signal ($t\bar{t}$) in red and background ($Z+$jets and $bb\Phi$ signals) in blue. The training and test samples are shown by the filled areas and points respectively. (b) The Receiver-Operator Characteristic curve for a selection based on the BDT score.

In order to use the BDT discriminant in an analysis selection, it is necessary to find the optimal value of the BDT score on which to cut. This can differ depending on whether signal acceptance or background rejection is more important, and this depends on the relative yields of signal and background. To study the potential analysis benefits of using this BDT, a search was performed to find the optimum threshold value and measure the enhancement of the sensitivity. Figure 5.4 shows the $s/\sqrt{b}$ - the number of signal events divided by the square root of the number of background events, a widely used approximation for the analysis sensitivity [196] - as a function of BDT score for each of the $bb\Phi$ signal samples. Only events with a dimuon mass within the Full Width at Half Maximum (FWHM) of the signal peak are considered. The peak for each signal represents the optimum choice of BDT score at which to cut and is normalised to unity and therefore the difference between the peak and plateau at BDT score ≈ 1 gives the improvement in sensitivity. This gain varies from a factor of 3.7 to approximately 1.5 for the $m_\Phi = 200$ GeV and $m_\Phi = 1$ TeV $bb\Phi$ signals respectively. The optimum choice of BDT score also varies as a function of signal mass, indicating that to truly benefit from the BDT, a different choice of cut would be needed for each signal hypothesis mass. Such a strategy would require further study, however, as it could introduce new systematic uncertainties into the analysis.

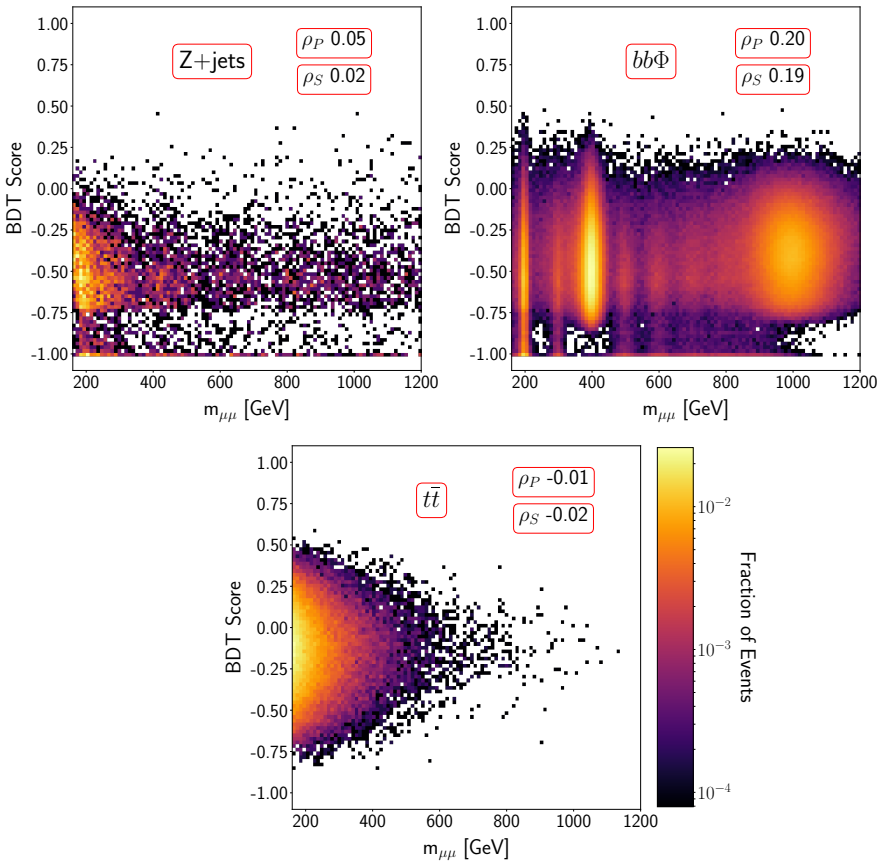


Figure 5.3: Heat-maps showing BDT score and dimuon invariant mass, $m_{\mu\mu}$, for the $Z+$jets (top left), $t\bar{t}$ (bottom) and all $bb\Phi$ signal samples (top right). Brighter areas contain a greater fraction of events while darker areas contain a small fraction of events. The Pearson and Spearman's rank correlation coefficients, ρ_P and ρ_S, are also shown [194, 195]. The $Z+$jets and $bb\Phi$ processes contain no true sources of E_T^{miss}, allowing good discrimination from $t\bar{t}$. Events with a BDT score of -1 typically have negligible reconstructed E_T^{miss}. The limited number of simulated $t\bar{t}$ events at high masses can also be seen, in contrast to the high-statistics $Z+$jets and $bb\Phi$ samples.

5.3.3 Summary

This section has outlined the selection requirements for physics objects to be included in the analysis. Furthermore, the categorisation of events into Signal and Control regions has been described. Finally, studies into the use of an MVA method to improve the sensitivity of the analysis has been explained. Whilst yielding promising results, this MVA method is not used for the full analysis described hereafter.

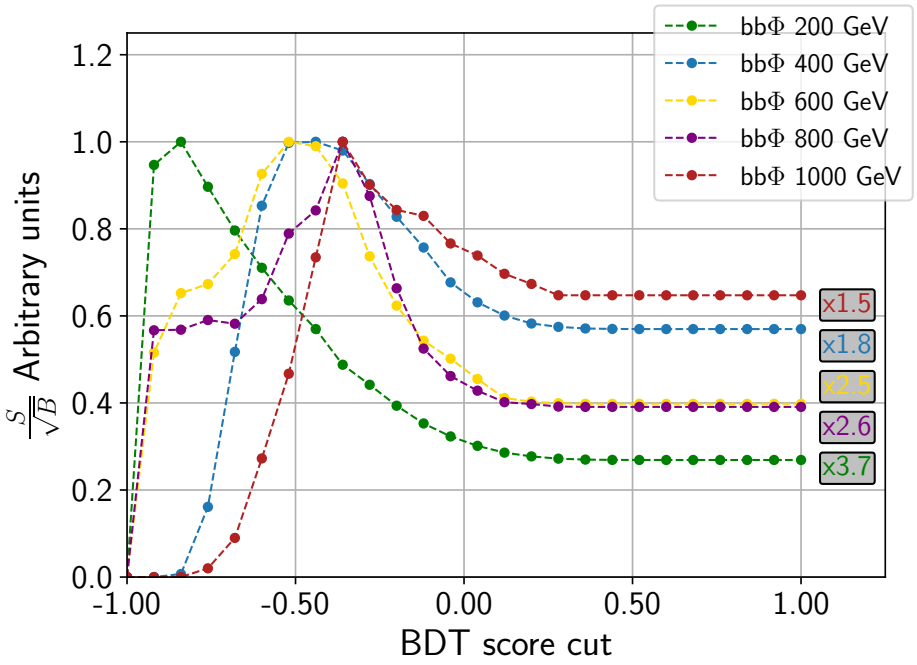


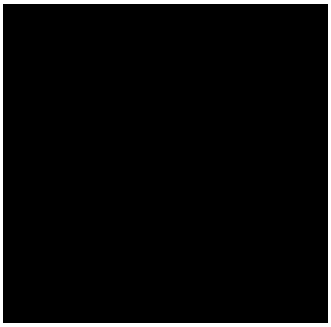
Figure 5.4: Change in $s/\sqrt{b}$, the signal sensitivity, for different thresholds of the BDT score. A requirement that the BDT score be less than 1 will always be true, so effectively represents no cut. Lower BDT score thresholds correspond to a tighter selection until a value of -1, when all events will fail the requirement. The displayed distributions for each signal mass are normalised such that the maximum sensitivity is unity. The maximum improvement in sensitivity is shown in grey boxes to the right and is calculated from the ratio of the maximum sensitivity to the sensitivity at the plateau at BDT score ≈ 1.

Chapter 6

Statistical Analysis

Never tell me the odds!

Han Solo



6.1 Likelihood function and test statistics

Cross-section times Branching Ratio ($\sigma_\Phi \times \mathcal{B}(\Phi \to \mu\mu)$) limits for the $bb\Phi$ and ggF production mechanisms are key results of this analysis. In order to extract limits on the $\sigma_\Phi \times \mathcal{B}(\Phi \to \mu\mu)$, a binned likelihood function is constructed from a product of Poissonian probabilities for each bin in the signal region

$$\mathcal{L}_{SR}(\mu, \theta) = \prod_{j \ \epsilon \text{ bins}} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)}, \tag{6.1}$$

where n_j are the number of observed events in each bin; s_j and b_j are the expected number of events which depend on so-called Nuisance Parameters (NPs) denoted by θ and μ is the signal strength parameter. This signal strength parameter is defined

as the ratio of the measured or excluded signal cross-section to the cross-section calculated by the MC generators described in Section 4.2.

In order to constrain systematic uncertainties and determine the normalisations of different background processes without signal contamination, measurements can be made in regions which are expected to contain no signal events. A likelihood function for a Control Region (CR) is constructed like so:

$$\mathcal{L}_{CR}(\theta) = \prod_{k \,\epsilon\, \text{bins}} \frac{b_k^{m_k}}{m_k!} e^{-b_k}, \tag{6.2}$$

where b_k and m_k are the expected and observed number of events in each bin of the control region. The total likelihood is simply the product of likelihoods from each signal and control region and gives a measure of the plausibility that these parameters $\mu,\ \theta$ have resulted in the observed data.

The Neyman-Pearson lemma [197] states that to test a hypothesised value of μ, the most powerful test is the profile likelihood ratio

$$\lambda(\mu) = \frac{L\left(\mu, \hat{\hat{\theta}}(\mu)\right)}{L\left(\hat{\mu}, \hat{\theta}(\mu)\right)}, \tag{6.3}$$

where $\hat{\hat{\theta}}$ are the values of the NPs that maximise the likelihood for a given μ whereas in the denominator $\hat{\mu}$ and $\hat{\theta}$ are the *unconditional* maximum-likelihood estimators. Since the denominator can vary over the full parameter space, the ratio is bounded between 0 and 1. Often the Log-Likelihood Ratio (LLR) is used as the natural log is monotonic, the maximum occurs for the same parameter values as the likelihood itself, and the likelihood is a very small number which would introduce larger floating-point errors than when using the log-likelihood.

There are many choices of test statistic that make use of the LLR and are widely used in the HEP community [196]. This analysis uses two, q_0 and $\tilde{q}_\mu$, for the discovery of a positive signal and upper limits respectively:

$$q_0 = \begin{cases} -2\log \frac{\mathcal{L}(\mu, \hat{\hat{\theta}}(\mu))}{\mathcal{L}(0, \hat{\hat{\theta}}(0))} & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases}, \tag{6.4}$$

$$\tilde{q}_\mu = \begin{cases} -2\log \frac{\mathcal{L}(\mu, \hat{\hat{\theta}}(\mu))}{\mathcal{L}(0, \hat{\hat{\theta}}(0))} & \hat{\mu} < 0 \\ -2\log \frac{\mathcal{L}(\mu, \hat{\hat{\theta}}(\mu))}{\mathcal{L}(\mu, \hat{\theta})} & 0 \leq \hat{\mu} \leq \mu \\ 0 & \mu < \hat{\mu} \end{cases}. \tag{6.5}$$

The reason for the difference between the two test statistics is that searching for a signal is subtly different to finding an upper limit on the possible signal strength. The q_0 test statistic quantifies whether the null, background-only hypothesis ($\mu = 0$) can be rejected. In this case, it is not desirable for a downward fluctuation in the data - such that $\hat{\mu}$ is negative - to count as a discrepancy since one expects the signal only to increase the number of events.

For the $\tilde{q}_\mu$ test statistic, used for finding upper limits, the opposite is true. One does not want to reject a given value of μ if the data fluctuates upwards and gives a best fit value of $\hat{\mu} > \mu$. Also, if the observed data gives $\hat{\mu} < 0$, the best level of agreement is for $\mu = 0$ as the signal strength can only physically be positive ($\mu \geq 0$).

The p-value is found from the pdf of the $\tilde{q_\mu}$ test statistic (with a similar definition of the p_0) as

$$p_\mu = \int_{\tilde{q}_{\mu,obs}}^{\infty} f(\tilde{q}_\mu|\mu) d\tilde{q}_\mu. \tag{6.6}$$

This pdf, $f(\tilde{q}_\mu|\mu)$, can be found through the use of many 'toy' experiments calculated by sampling all of the NPs to create an ensemble that spans the entire parameter space. The pdf is simply the distribution of the test statistic calculated for each toy. Alternatively a set of closed-form, asymptotic formulae can be used in the large data sample limit [196]. An example of a test statistic distribution found using each method is shown in Figure 6.1b.

When searching for a signal, the p_0 is used to quantify the significance of any excesses in the observed data. To make it more readable, the p_0 can be converted into a significance via $Z = \Phi^{-1}(1-p)$ where Φ is the cumulative distribution of the standard (zero mean, unit variance) Gaussian. If the significance reaches the thresholds - arbitrarily decided but nonetheless widely accepted - of 3σ or 5σ (corresponding to p_0 values of 1×10^{-3} and 3×10^{-7}), one can respectively claim "evidence for" or "discovery of" new physics.

Rather than determining the upper-limits on $\sigma_\Phi \times \mathcal{B}(\Phi \to \mu\mu)$ from the p-value of the signal-plus-background hypothesis, p_{s+b}, they are found using the CL_s method with the following definition [198]:

$$CL_s \equiv \frac{p_{s+b}}{1 - p_b} < \alpha. \tag{6.7}$$

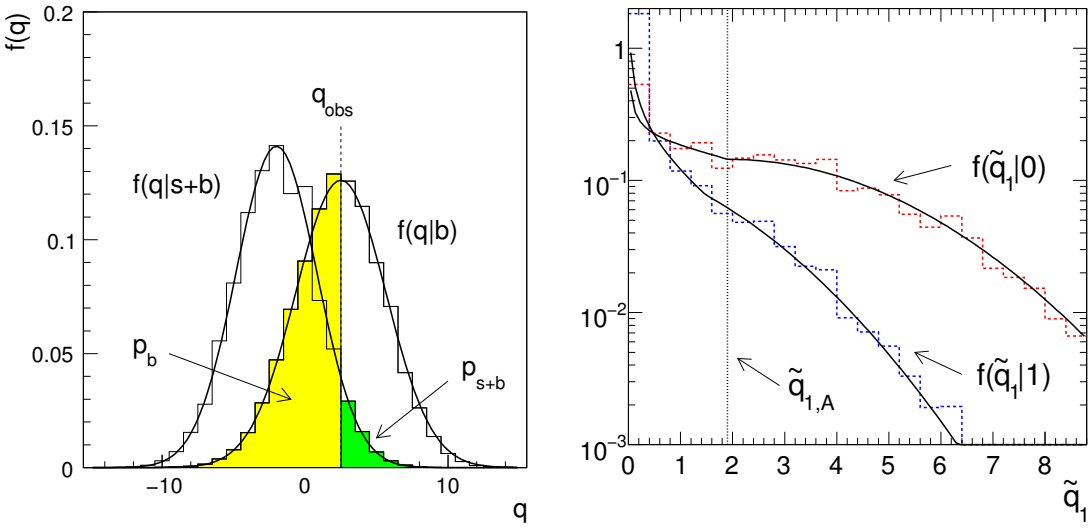
The reasoning behind using the CL_s method is that one expects the requirement $p_{s+b} < \alpha$ to exclude, with a probability close to α, hypotheses to which the analysis is not sensitive. Consider the example distributions shown in Figure 6.1a, where

there is overlap between the $s+b$ and b only cases. If the analysis is not sensitive to the signal, then the distributions overlap perfectly. The calculated p-value will then meet the requirement regardless of signal strength. To counteract this, the p_{s+b} is penalised by the divisor $1-p_b$ where p_b is defined to be:

$$p_b = \int_{-\infty}^{\tilde{q}_{obs}} f(\tilde{q}_\mu|\mu) d\tilde{q}_\mu. \tag{6.8}$$

If the distributions $f(\tilde{q}_\mu|s+b)$ and $f(\tilde{q}_\mu|b)$ are well separated, $1-p_b$ is only slightly less than unity, and the penalty is small. However, the penalty becomes larger if the distributions overlap, preventing the exclusion of signal models to which the analysis is not sensitive.

The 95% Confidence Level (CL) corresponds to the number of signal events for which the α is 0.05 and is a misnomer, or at the least misleading, as the CL_s is not a true confidence interval since the CL of the interval is not equal to the coverage probability. Just as for the pdf of the test statistic, the median and error bands for the upper limit can either be determined using toy experiments or an asymptotic approximation.



(a) (b)

Figure 6.1: The CL_s method (a) compares the test statistic for the background only, b and signal-plus-background $s+b$ hypotheses for some generic test statistic [198], q. This analysis uses the $\tilde{q}_\mu$ test statistic (b). The red and blue histograms show examples of this distribution for a background only and signal-plus-background hypothesis respectively, found using many 'toy' experiments. The asymptotic approximations are shown as grey curves and match the toy distributions well [196].

6.1.1 Nuisance Parameters

The expected signal and background event yields, s_j and b_j in eqn. (6.1), are dependent on a large number of systematic uncertainties, some from experimental measurements and others from theoretical predictions. These systematic uncertainties are outlined in more detail later in Chapter 7 and their impact on the final results of the analysis is achieved through the use of Nuisance Parameters (NPs). The term *nuisance* is used as the parameter itself is not of immediate interest but it must be accounted for nonetheless. These NPs describe how far each systematic effect is from its nominal value. For instance, the muon momentum scale NP encodes the degree to which ATLAS is systematically under- or over-estimating the true momentum of a muon. Each of the NPs that describes systematic uncertainties enters the likelihood function not only indirectly through its effect on the event yields, s_j and b_j but also through a Gaussian constraint term:

$$\mathcal{G}_{\text{NP constraint}}(\hat{\theta}|\theta_0, \sigma) = e^{-\frac{(\hat{\theta}-\theta_0)}{2\sigma^2}}, \tag{6.9}$$

where $\hat{\theta}$ is the NP, θ_0 is the NP nominal value, and σ is the size of uncertainty. When included in the likelihood, these terms constrain the NP to remain close to the nominal value as the Gaussian term penalizes the likelihood the further the NP is from the nominal. The size of the uncertainty, σ, determines how tightly the NP should be constrained. Much of the following chapter concerns the steps taken to estimate the sizes of the uncertainties so they can be included in the likelihood fit.

Before moving on from the topic of nuisance parameters, it should also be noted that the normalisations of the major backgrounds are also included as NPs in the likelihood fit as they impact the expected background yield, b_j. However, contrary to the systematic uncertainty NPs, the normalisation factors are 'freely-floating' parameters meaning that no Gaussian penalty term is applied to the likelihood.

6.2 Analysis Fit Strategy

The likelihood function used in this analysis is constructed from binned distributions of the dimuon invariant mass in the two Signal Regions (SRs) and three Control Regions (CRs). In each of the CRs there is only one bin, used to measure the normalisations of the dominant background processes. In the SRs, there are 61 bins with each bin increasing exponentially in width. This *log-uniform* binning allows for fine resolution at low masses - where the experimental resolution is also small -

whilst ensuring that there are enough events in each of the bins at high mass, where the expected event yield is lower.

The binned dimuon invariant mass distributions are stored in so-called *template* histograms which are produced for each background and signal process. These histograms are used as inputs to the HistFactory package [199] which constructs the statistical model within the RooFit framework [3]. Template histograms describing each source of systematic uncertainty are also included as inputs to describe $\pm 1\sigma$ variations of the Nuisance Parameters (NPs) and are interpolated linearly. The RooFit package is then used to perform two types of fit: searching for signal using the q_0 test statistic, eqn. (6.4), and placing upper limits on the cross-section with the $\tilde{q}_\mu$ test statistic, eqn. (6.5). In both cases, the Parameter of Interest (PoI) is the signal strength, μ. The upper limit on the cross-section can be found by multiplying the upper limit on the signal strength by the cross-section calculated when generating the relevant signal MC sample.

6.2.1 Asimov Dataset

It is often useful when understanding the results of search and exclusion fits, to compare results of a fit using the observed data to one using a so-called *Asimov dataset* [196]. An Asimov dataset is effectively a single MC toy, where the values of every parameter are set equal to their expected values. By comparing this 'nominal' dataset with the observed data, one can more easily understand how any experimental measurements have differed from one's expectations. An Asimov dataset can also be used to determine the expected sensitivity of an analysis, as well as the impact the variations of different NPs have on the Parameter of Interest (PoI). The RooFit package [3] was used to generate an Asimov dataset which is used to compare to the observed data results, shown later in Chapter 8.

6.3 Signal Interpolation

The signal MC samples described in Section 4.2 can be used to test signal hypotheses at the nine mass points for which samples were generated. However, as is shown in Figure 6.2, the muon momentum resolution is only a few percent for muon p_{T} values of approximately 100 GeV. This means that dimuon invariant mass peaks at the lower end of the 0.2-1 TeV search range will be far smaller than the 100 GeV spacing between signal samples. To search for these peaks at intermediate mass points and to prevent generating an excessive number of signal samples, a signal interpolation procedure was developed.

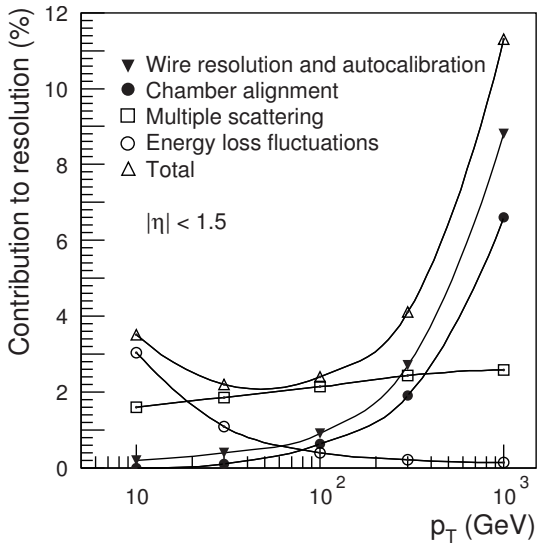


Figure 6.2: Contributions to muon momentum resolution as functions of p_{T} as expected from Monte Carlo simulations. The overall momentum resolution increases significantly between 100 GeV and 1 TeV [200].

The Double-sided Crystal Ball (DSCB) function [201] was found to fit the dimuon mass shape of the signal samples at all masses and is defined as

$$DSCB(M_{\mu\mu};\bar{x},\sigma,\alpha_L,\alpha_H,n_L,n_H) = \begin{cases} e^{-(\frac{(M_{\mu\mu}-\bar{x})^2}{2\sigma^2})} & \text{for } \alpha_L < \frac{M_{\mu\mu}-\bar{x}}{\sigma} < \alpha_H \\ (\frac{n_L}{|\alpha_L|})^{n_L} \times (\frac{n_L}{|\alpha_L|} - |\alpha_L| - \frac{M_{\mu\mu}-\bar{x}}{\sigma})^{-n_L} & \text{for } \alpha_L \geqslant \frac{M_{\mu\mu}-\bar{x}}{\sigma} \\ (\frac{n_H}{|\alpha_H|})^{n_H} \times (\frac{n_H}{|\alpha_H|} - |\alpha_H| - \frac{M_{\mu\mu}-\bar{x}}{\sigma})^{-n_H} & \text{for } \alpha_H \leqslant \frac{M_{\mu\mu}-\bar{x}}{\sigma} \end{cases} . \tag{6.10}$$

It has a Gaussian core with power-law tails and has six parameters: the mean $\bar{x}$ and standard deviation σ and two parameters for each tail α_L, n_L and α_H, n_H for the low and high-mass tails respectively.

Each of these six parameters is expected to change with the mass of the hypothetical signal resonance, therefore they are each parameterised themselves with second-order polynomials (see Table 6.1). An example of the change in DSCB parameters with mass is shown in Figure 6.3.

To interpolate between the differences in selection acceptances for different signal masses, each signal sample was reweighted from its generated cross-section to the same, arbitrary cross-section of 10 nb. The differences in the event yield of each sample are then due only to changes in the acceptance. The normalisation of the DSCB is also included as a parameter in the signal interpolation and is similarly parameterised itself with a second order polynomial.

Table 6.1: Parameterisations of DSCB parameters as functions of signal hypothesis mass m_Φ. Cases where the higher-order terms were negligible have been removed.

DSCB Parameter	**Polynomial**
Normalisation	$a_{Norm} + b_{Norm} m_\Phi + c_{Norm} m_\Phi^2$
Mean ($\bar{x}$)	$a_{\bar{x}} + b_{\bar{x}} m_\Phi$
Width (σ)	$a_\sigma + b_\sigma m_\Phi + c_\sigma m_\Phi^2$
α_L	$a_{\alpha_L} + b_{\alpha_L} m_\Phi$
n_L	$a_{n_L} + b_{n_L} m_\Phi$
α_H	$a_{\alpha_H} + b_{\alpha_H} m_\Phi$
n_H	a_{n_H}

After fitting to the signal samples to determine the polynomial coefficients, the final signal shape is only a function of the hypothesis mass, m_Φ. It is then possible to produce signal templates to be used in the search and exclusion fits described in the previous section. The fit and interpolation is performed separately for signal MC samples corresponding to every $\pm 1\sigma$ variation of the NPs described in Chapter 7.

This interpolation process was validated by removing each of the nine signal samples separately and then comparing the interpolated DSCB shape directly with the removed MC template mass shape. In all cases the reduced χ^2 [202], a measure of goodness-of-fit, was less than 1.35 although it was noted that the agreement was worst for the lowest and highest mass points, 200 GeV and 1 TeV. This suggests that the somewhat simplistic approach used here is not suitable for extrapolating the invariant mass shape to lower or higher values and so it is only used to interpolate to intermediate mass points. For this reason, the lowest hypothesis mass considered in this analysis is $m_\Phi = 200$ GeV even though the search region extends to a dimuon invariant mass of $m_{\mu\mu} = 160$ GeV.

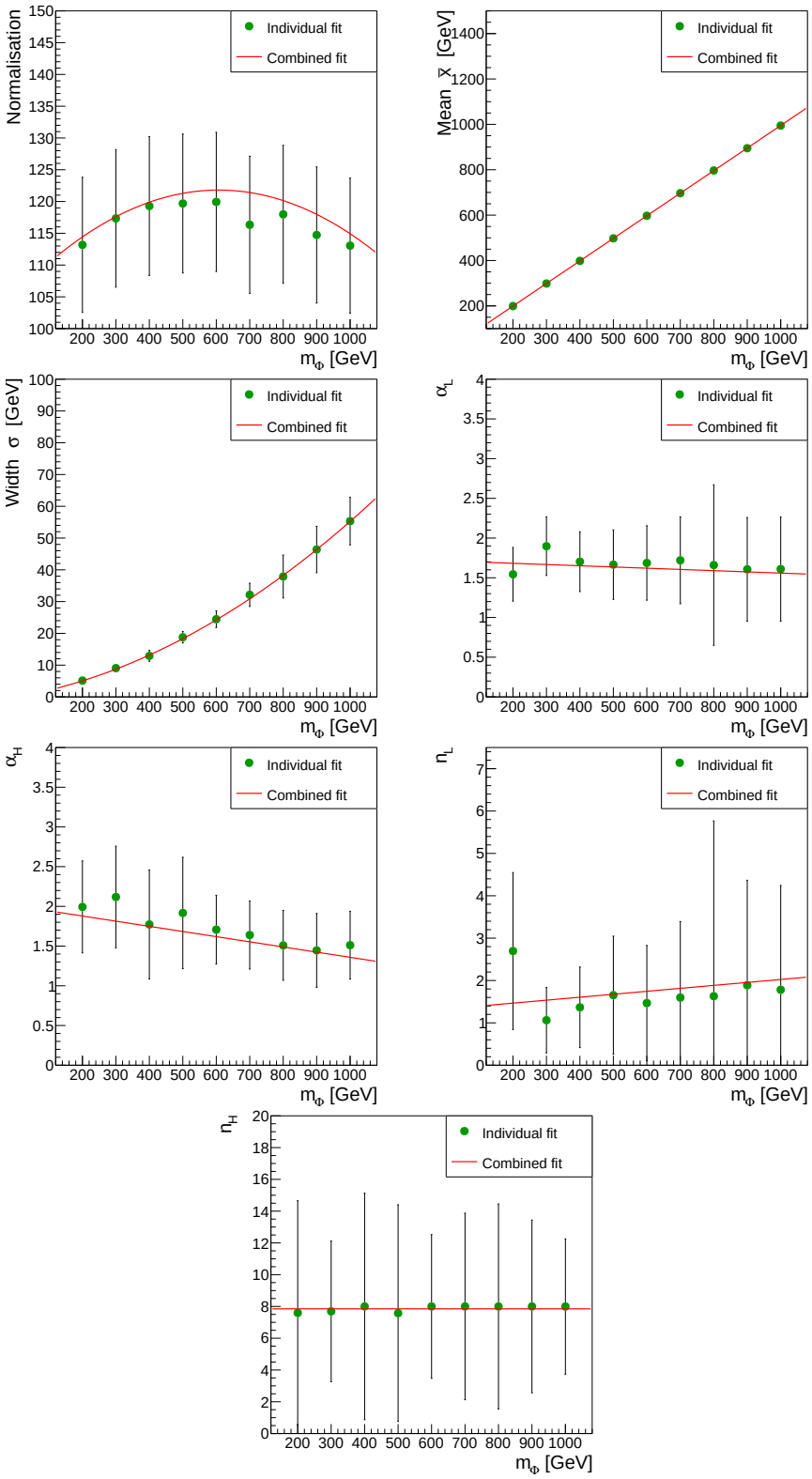


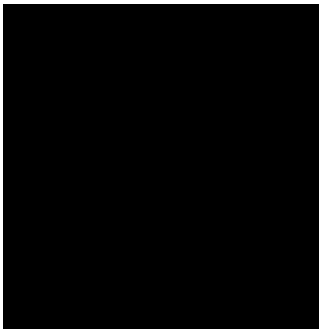
Figure 6.3: Change in DSCB parameters with signal hypothesis mass m_Φ for the ggF signal in the b-Veto signal region. The green points show the fitted parameter for a fit of each mass separately and the red curve shows the fitted polynomial that results from the combined fit of all samples.

Chapter 7

Systematic Uncertainties

It is far better to foresee even without certainty than not to foresee at all.

Henri Poincaré



The objective of any search for Beyond Standard Model physics is to compare measured data with expectations, looking for discrepancies. To properly quantify the significance of any results it is necessary to take into account the systematic uncertainties that will affect both the measurements and predictions. This chapter discusses each of the sources of systematic uncertainty considered, including the methods used to estimate the size of the uncertainty. The first of the sections contains details on the experimental sources of uncertainty and the latter concerns the theoretical uncertainties.

7.1 Experimental sources of uncertainty

This analysis makes use of the reconstructed physics objects, described in Chapter 3, which are themselves composite measurements made by each ATLAS sub-detector.

It is the task of the so-called *Combined Performance* groups within the ATLAS collaboration to calibrate these physics objects and estimate the relevant uncertainties. This section begins with the most important experimental sources of uncertainty; involving muons, jets and the b-Tagging algorithms and then other, miscellaneous uncertainties are discussed.

7.1.1 Muon measurements

As the experimental signature of this analysis contains two muons, it is critical that they are identified and reconstructed well. Muons are reconstructed and calibrated by following the procedure as described in Section 3.5. The measurements and corrections have uncertainties which must be taken into account.

Firstly, scale factors that correct for the difference in efficiencies between MC and data when reconstructing and triggering on muons are measured in *tag-and-probe* studies, looking at $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ events. The largest sources of uncertainty in determining these scale factors are correlations between the 'tag' and 'probe' muons, contributions from background processes, and statistical uncertainties in the data and MC samples [122]. The total effect of these uncertainties on the scale factors are found to be approximately 1% in the p_{T} region studied. An upper limit on the variation at higher p_{T} is extracted from simulation and so an additional uncertainty is applied of 2-3% per TeV for muons with $p_{\mathrm{T}} > 200$ GeV. Similar scale factors are derived for the efficiencies of the isolation and Track-to-Vertex Association (TTVA) selections. In each case, the component of the uncertainty that is statistical is separated from the other sources of uncertainty.

Measurements of the muon momentum are calibrated using fits to the dimuon invariant mass peaks of the Z boson and J/ψ meson and uncertainties on this calibration are determined from the error on the fitted parameters. The uncertainties on the momentum scale are found in bins of muon pseudorapidity and momentum and vary from 0.05% in the barrel region to 0.1% at $|\eta| \approx 2.5$. For the muon resolution, the uncertainties vary from 2.3% in the central region to 2.9% in the endcaps. A complete list of the muon measurement uncertainties can be found in Table B1 in Appendix B.

7.1.2 Jet measurements

There are many steps in reconstructing, measuring and calibrating jets in ATLAS, described in Section 3.3 and therefore there are many sources of uncertainty that one must consider. An eigenvector variation method is used to reduce the number

of nuisance parameters that will be used in the fit [203]. Firstly the covariance matrix of all possible variations is found and, as it is a symmetric positive-definite matrix, it is trivial to find the eigenvectors and eigenvalues of this matrix. These eigenvectors can be thought of as 'directions' in which variations are independent from one another. If the effect of varying along an eigenvector is negligible it can be removed, reducing the total number of parameters that must be included. In this way, the 88 total NPs related to jet measurements can be decomposed into 21 uncorrelated NPs and each is estimated in bins of jet p_{T} and η. These uncertainties are listed in Table B2 in Appendix B.

7.1.3 Flavour Tagging uncertainties

A correction factor is applied to the outcome of the flavour tagging algorithm described in Section 3.4 to account for differences in the efficiency between data and simulation. This flavour tagging correction factor is derived separately for b-jets, c-jets and light-flavour jets, and just as for the jet energy scale and resolution correction factors, it is dependent on the jet p_{T} and $|\eta|$ and is affected by uncertainties from multiple sources. These sources can be broadly categorised into three groups: MC generator modelling of kinematic variables; normalisation uncertainties of the cross-section of simulated samples and experimental uncertainties arising from detector and reconstruction effects. An eigenvector variation method [203] is used to reduce the number of nuisance parameters, resulting in three NPs for b-jets and for c-jets and five NPs for light-flavour jets. The approximate size of the uncertainty in the flavour tagging efficiency is 2% for b-jets, 10% for c-jets and 30% for light-flavour jets [204]. These uncertainties were estimated in studies of $t\bar{t}$ events where the maximum jet p_{T} was approximately 300 GeV, therefore an additional uncertainty is applied to the extrapolation of the flavour tagging efficiency for jets with $p_{\mathrm{T}} > 300$ GeV. These uncertainties are listed in Table B3 in Appendix B.

7.1.4 Other Experimental Uncertainties

The measured integrated luminosity of the dataset has an uncertainty of 3.2%, the largest source of which is the calculation of the total visible cross-section performed during Van der Meer (VdM) scans [100]. These uncertainties come about from instrumental effects within the ATLAS luminometers, as well as changes in the exact beam conditions of the LHC.

As the E_T^{miss} is a topological event quantity, its uncertainties are mostly due to the measurement of objects used in its calculation. For instance, in the case of

muons and jets the uncertainties are found as described above. However, the 'soft' term - tracks and calorimeter deposits that have not been reconstructed as physics objects - also has an uncertainty which must be separately determined. In a event with no true source of missing energy, the soft term would be expected to perfectly cancel the hard term. For the purposes of measuring the soft term, it is projected along the direction antiparallel to the hard term and three quantities are measured: the energy scale of the soft term parallel to this axis, and the resolution of the transverse and parallel components. These quantities are compared in bins of p_T^{hard} between data and three MC generators and the maximal disagreement is taken as a systematic uncertainty [205].

7.2 Theoretical sources of uncertainty

There are many theoretical sources of uncertainty in the modelling of the background and signal processes. This section describes each of these sources and the methods used to quantify the size of each uncertainty. At first, the uncertainties related to the background processes will be described followed by a discussion of the uncertainties considered for the signal model.

7.2.1 Background modelling uncertainties

The background MC samples are used to predict the dimuon invariant mass ($m_{\mu\mu}$) distribution and therefore systematic uncertainties are evaluated with regards to their effects on the shape and normalisation of this distribution. The uncertainties applied to the two major background processes, $Z+$jets and $t\bar{t}$ production, are discussed in detail in the following two sections and have a larger impact than the uncertainties on the other background processes which are subsequently summarised.

$Z+$jets $m_{\mu\mu}$ Shape and Extrapolation systematics

Uncertainties on the $Z+$jets background samples are evaluated for the variation and choice of PDFs, the energy scale, the uncertainty on α_S, as well as the uncertainties on the higher order Electroweak (EW) and Photon-Induced (PI) corrections described in Section 4.3.1. Only the effect on the shape of the $m_{\mu\mu}$ is considered as the normalisations of the Z+LF and Z+HF components are included later, in the fit, as freely-floating parameters.

The PDF variation uncertainty was obtained using the 90% confidence level CT14NNLO PDF error set and by following the procedure described in Refs. [206,

207]. Rather than using a single Nuisance Parameter (NP) to describe the 28 eigenvectors of this PDF error set, which could lead to an underestimation of its effect, a re-diagonalized set of 7 PDF eigenvectors was used [161]. This represents the minimal set of PDF eigenvectors that maintains the necessary correlations, and the sum in quadrature of these eigenvectors matches the original CT14NNLO error envelope well. They are treated in the likelihood fits as separate NPs. The uncertainties due to the variation of PDF scales and α_S were derived using the VRAP program [174]. The former was obtained by varying the renormalisation and factorisation scales of the nominal CT14NNLO PDF up and down simultaneously by a factor of two. The value of α_S used (0.118) was varied by ± 0.002. The EW correction uncertainty was assessed by comparing the nominal additive $(1 + \delta_{\text{EW}} + \delta_{\text{QCD}})$ treatment with the multiplicative approximation $((1 + \delta_{\text{EW}})(1 + \delta_{\text{QCD}}))$ treatment of the EW correction in the combination of the higher-order EW and QCD effects. The uncertainty in the photon-induced correction was calculated from the uncertainties in the quark masses and the photon PDF. Following the recommendations of the PDF4LHC forum [207], an additional uncertainty due to the choice of nominal PDF set was derived by comparing the central values of CT14NNLO with those from other PDF sets, namely MMHT14 [208] and NNPDF3.0 [209]. The maximum width of the envelope of these comparisons was used as the PDF choice uncertainty, but only if it was larger than the width of the CT14NNLO PDF eigenvector variation envelope.

The discrepancies in the modelling of the Z+HF discussed in Section 4.4.1 motivate the use of an additional uncertainty on the $m_{\mu\mu}$ shape for this component of the background. The default Powheg Monte Carlo events are compared with the predictions of Sherpa 2.2.1 [181] and MadGraph 5 [159] which implement more sophisticated modelling of the Z+HF jets. The difference of the Powheg prediction to the other generators as a function of $m_{\mu\mu}$ discriminant is used to motivate the shape of an additional uncertainty on the Z+HF prediction in the b-Tag Signal Region with a functional form $\pm$ 0.4 $\ln(\frac{m_{\mu\mu}}{300\ \text{GeV}})$ as shown in Figure 7.1.

The normalisations of the $Z+$jets backgrounds are predominantly determined from the large number of events in the Control Regions (CRs) with only a small contribution from the Signal Regions (SRs). When considering the expected yield in the Signal Regions, additional uncertainties are placed on this extrapolation of the normalsiation factor from the CRs. The ratio between the expected SR and CR yields are calculated for $\pm 1\sigma$ variations of each source of uncertainty listed above. For the Z+HF component, the ratio is also calculated for alternative choices in Monte Carlo event generator. The fractional differences of the SR:CR ratio with respect to the nominal are added in quadrature and this total is taken

as the overall uncertainty on the extrapolation. These uncertainties are labelled as `Ratio_ZjetsLF_CRSR` and `Ratio_ZjetsHF_CRSR` with magnitudes of 2% and 7% for the light-flavour and heavy-flavour components respectively.

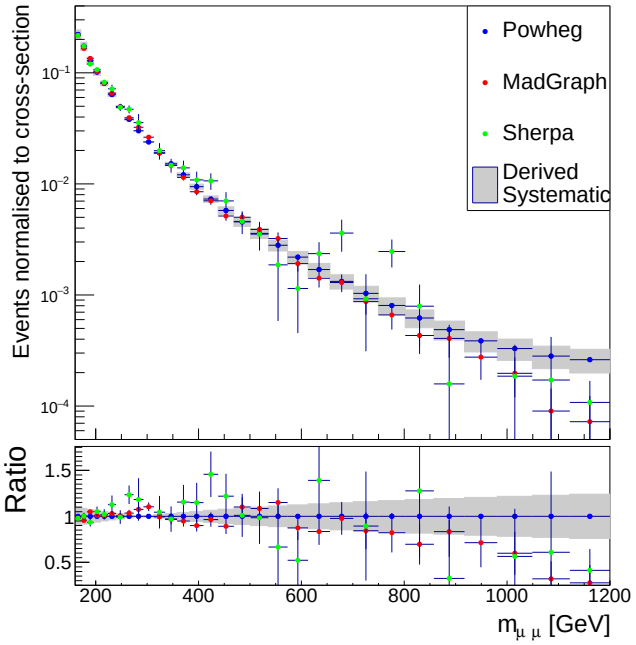


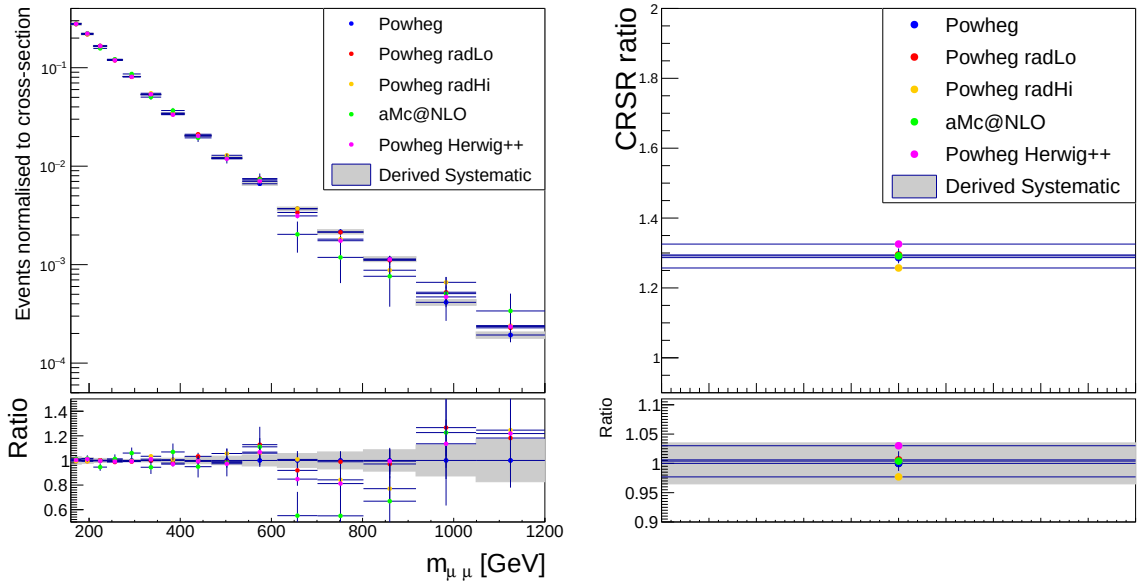
Figure 7.1: $m_{\mu\mu}$ shape predicted by the POWHEG, MADGRAPH and SHERPA generators. The shaded band shows the derived shape uncertainty.

$t\bar{t}$ $m_{\mu\mu}$ Shape and Extrapolation Systematics

An $m_{\mu\mu}$ shape uncertainty for the $t\bar{t}$ background is derived in an analogous way to the $Z+$jets uncertainty described previously. To estimate the impact of Initial State Radiation (ISR) and FSR, two alternative POWHEG + PYTHIA samples were generated with modified parameters. In one case, the renormalisation (μ_R) and factorisation (μ_F) scales were halved whilst the h_{damp} parameter of POWHEG, which controls parton shower matching and effectively regulates the amount of high-p_{T} radiation, was doubled to twice the mass of the top-quark. In the other sample, the renormalisation and factorisation scales were doubled and the h_{damp} parameter was unchanged. In both cases the corresponding Perugia 2012 radiation tune was used. These choices of parameters have been shown to encompass the cases where μ_R and μ_F are varied independently and covered the measured uncertainties of the data in unfolded $t\bar{t}$ distributions [210, 211].

As explained in Section 4.1.2, there are several choices for parton shower and hadronisation models. Differences in the $t\bar{t}$ $m_{\mu\mu}$ shape for different choices of these models were studied by interfacing POWHEG-BOX with HERWIG++ 2.7.1 [148] with the UE-EE-5 tune [162] and the corresponding CTEQ6L1 PDF sets. Finally, a sample was generated with MADGRAPH5_AMC@NLO 2.2.1 [159] also interfaced to HERWIG++. An NLO matrix element and CT10 PDF were used for the $t\bar{t}$ hard scattering process, and HERWIG++ 2.7.1 modelled the parton shower, hadronisation and underlying event.

A functional form was chosen which encompassed differences in the $m_{\mu\mu}$ shape between the nominal sample and these alternative samples. The variation, labelled `TTBAR_MLLSHAPE` has the form: $\pm$ 0.2 $\ln(\frac{m_{\mu\mu}}{300\ \text{GeV}})$. Just as for the Z + jets extrapolation uncertainties above, differences in the ratio of the expected yields of the signal and control regions were used to estimate a 3.5% uncertainty on the extrapolation from CRttbar to SRbTag, labelled `Ratio_ttbar_CRSR`. The $m_{\mu\mu}$ shape and extrapolation uncertainties are shown in Figure 7.2.



(a) (b)

Figure 7.2: Comparison of (a) $m_{\mu\mu}$ shape and (b) control-signal region event yield ratios for different MC generators.

Other background Systematics

Due to the very small contribution of the Single-Top and Diboson background processes to the event yield in the signal regions, no uncertainty on the $m_{\mu\mu}$ shape is considered. Uncertainties arising from higher-order contributions to the cross-section and the choice and variation of PDFs are evaluated in a similar manner to the major backgrounds, however only the effect on the overall normalisation is considered. The resulting normalisation uncertainties, labelled `XS_SingleTop` and `XS_Diboson`, have magnitudes of 5.4% and 6% respectively.

7.2.2 Signal theory systematics

Sources of uncertainty that are related to the signal modelling include initial and final state radiation, the modelling of Multi-parton Interaction (MPI), the renormalisation and factorisation scale, and finally the Parton Distribution Functions (PDFs).

In order to estimate the impact on the uncertainties of factorisation and renormalisation scales, they were varied by a factor of 2 up and down, including correlated and anti-correlated variations. For the $bb\Phi$ process, the largest deviation from the nominal value of the acceptance was taken as a conservative estimate of the final scale uncertainty. For the ggF process a 25% uncertainty on the acceptance in the SRbTag region is considered as well as the anti-correlated effect in the SRbVeto region, following the Stewart-Tackman procedure adopted in [212].

The estimation of the uncertainty due to the PDF was performed by reweighting the signal samples for every PDF in the `PDF4LHC15_nlo_100` and `PDF4LHC15_nlo_nf4_30` sets for $bb\Phi$ and ggF signals respectively [207]. The envelope of the resulting variations in acceptance was taken as the combined PDF uncertainty.

The uncertainties arising due to ISR, FSR and MPI modelling were derived through studies of the acceptance change for variations in the parton shower tune. The total uncertainty is defined as the quadratic sum of the resulting fractional change in acceptance for each of the tune variations.

All of these uncertainties were evaluated for samples at 3 mass points: 200, 400 and 1000 GeV. As the variations are small and for the sake of simplicity, the largest of the three estimated uncertainties are used for all mass points. None of the considered generator parameter variations resulted in a statistically significant effect on the shape of the $m_{\mu\mu}$ distribution so all systematic effects are considered only as normalisation uncertainties. The name and size of the theory uncertainties applied to signal samples are shown in Table 7.1. The uncertainties associated with the same nuisance parameter are fully correlated and the sign of the uncertainty

shows whether the effect on the signal acceptance is correlated or anti-correlated. For instance, an increase in the QCD scale leads to an increase in the acceptance of $bb\Phi$ events in both signal regions, whereas for the ggF process the same change leads to an anti-correlated change between the b-Tag and b-Veto signal regions.

Table 7.1: The largest percentage change in the acceptance of the ggF and $bb\Phi$ for three different mass points (200, 400, 1000 GeV), due to variations of factorisation and renormalisation scales (`QCDscale`), variations in parton shower tune (`Tune`) and the PDF uncertainties (`PDF`). The NP name also includes to which signal model and region the uncertainty is relevant.

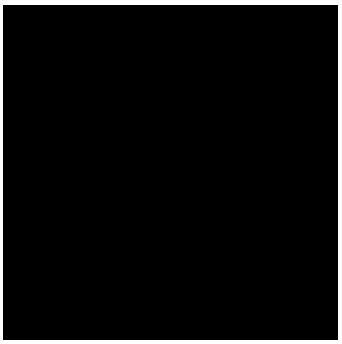
	bbΦ		***ggF***	
NP name	**SRbTag**	**SRBVeto**	**SRbTag**	**SRBVeto**
QCDscale_bbPhi	+1.8%	+1.1%		
QCDscale_ggF_1jet			+25%	-0.13%
Tune_bbPhi	+3.6%	-3.2%		
Tune_ggF			+3.8 %	-0.4%
PDF_bbPhi_bTag	0.87%			
PDF_bbPhi_bVeto		0.73%		
PDF_ggF_bTag			1%	
PDF_ggF_bVeto				1%

Chapter 8

Results

I love it when a plan comes together!

John "Hannibal" Smith



8.1 Post-fit mass distributions

The post-fit event yields in each of the Control Regions (CRs) are shown in Figure 8.1, where there is only one bin in the range 100 GeV $< m_{\mu\mu} <$ 160 GeV. The CRbVeto and CRttbar regions have a high purity of the *Z*+*LF* and $t\bar{t}$ backgrounds respectively, whereas the CRbTag region contains appreciable contributions from each of the major backgrounds. The normalisation factors of the backgrounds resulting from the fit are listed in Table 8.1. One can see that the $t\bar{t}$ and *Z*+*LF* backgrounds are scaled up by only 2%, whereas the *Z*+*HF* background component is increased by 46%, as expected from the modelling validation described in Section 4.4.

The post-fit $m_{\mu\mu}$ distributions for the Signal Regions are shown in Figure 8.2. There is very good agreement between the data and expected background in the bVeto SR. In the bTag region, with fewer events expected in each bin, the statistical fluctuations are expected to be higher. Indeed, the ratio between data and expected background varies considerably between some bins in the high mass region, although generally there is still reasonably good agreement within statistical and systematic uncertainties.

Table 8.1: Post-fit normalisation factors.

Normalisation Factor	Asimov dataset	Observed dataset
`mu_Zjets_HF`	1.00 ± 0.23	1.46 ± 0.26
`mu_Zjets_LF`	1.00 ± 0.02	1.02 ± 0.02
`mu_ttbar`	1.00 ± 0.04	1.02 ± 0.04

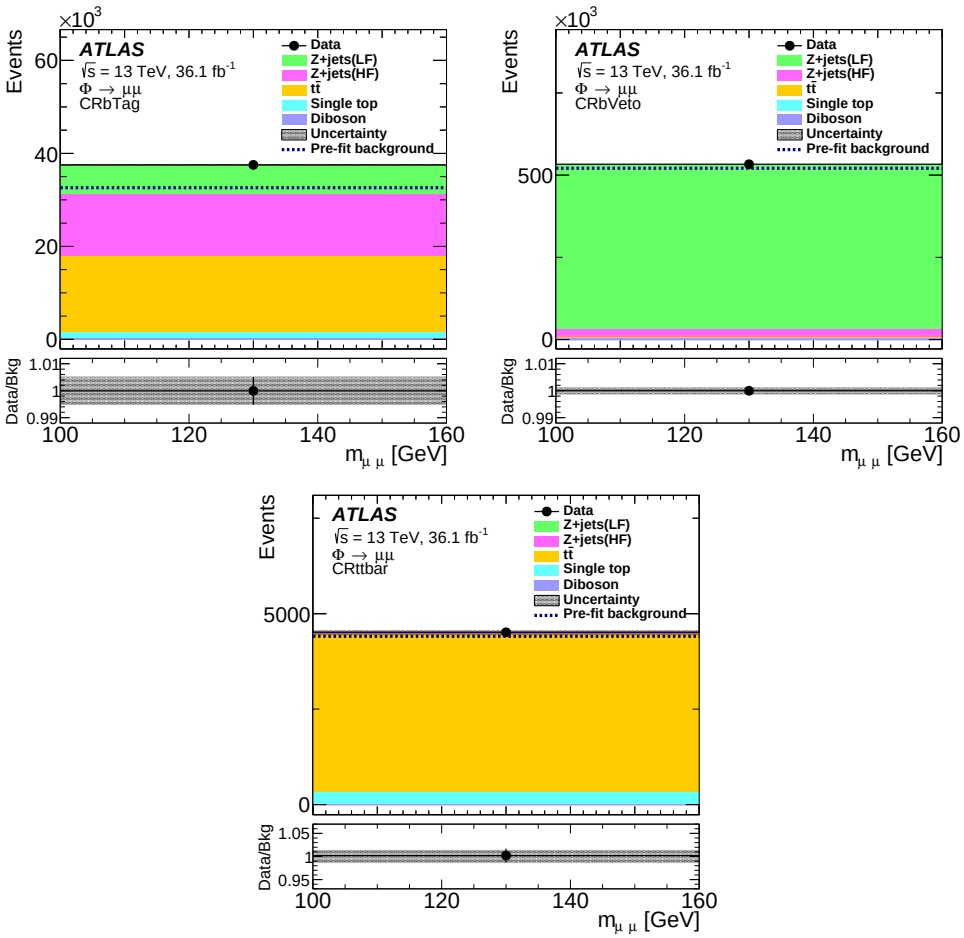


Figure 8.1: Post-fit event yields in each of the control regions. Each background process is normalised according to its post-fit cross-section. The data are shown by the points, while the size of the statistical uncertainty is shown by the error bars. The hatched band shows the total systematic uncertainty of the post-fit yield.

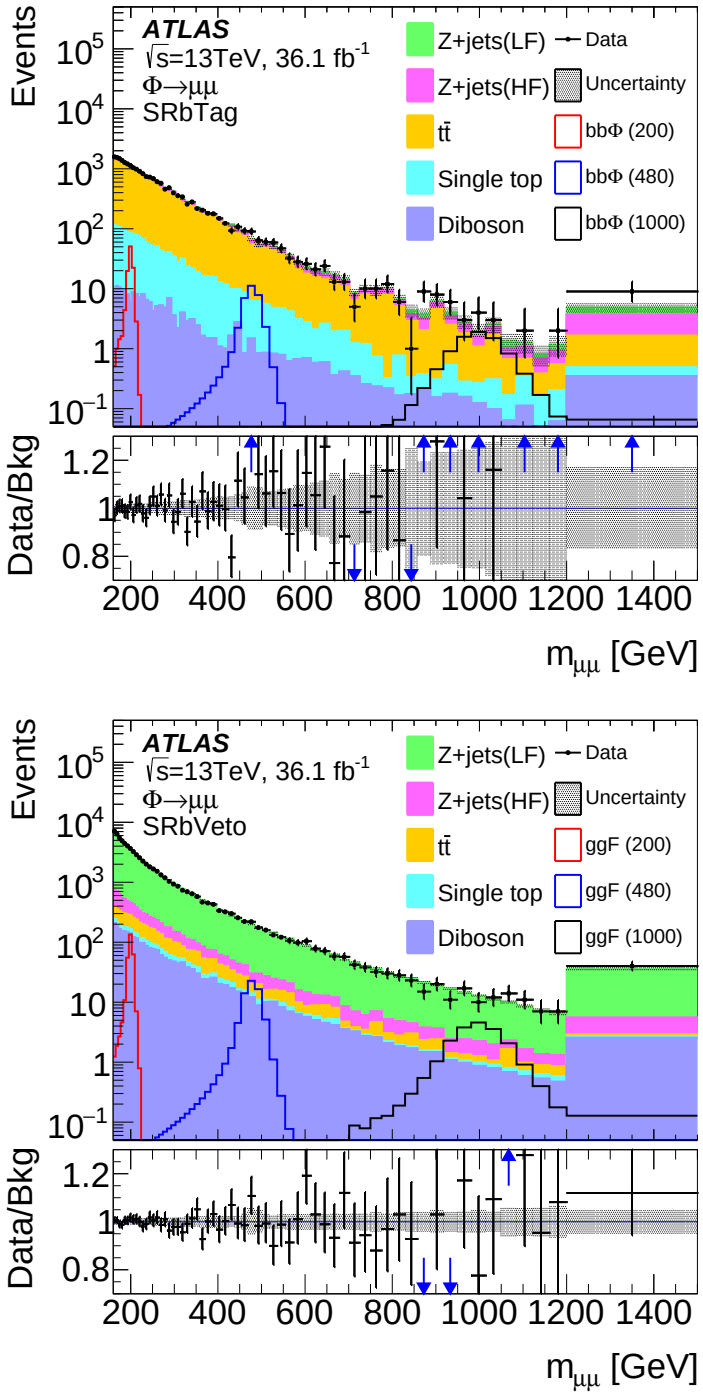


Figure 8.2: Distributions of the dimuon invariant mass, $m_{\mu\mu}$, after the combined fit to data under the background-plus-signal hypothesis ($bb\Phi$, $m_\Phi = 480$ GeV in SRbTag and ggF, $m_\Phi = 480$ GeV in SRbVeto). The fit for a signal corresponds to the largest excess observed above the background expectation. Each background process is normalised according to its post-fit cross-section. The templates for the signal mass hypotheses are normalised to their expected upper limits. The data are shown by the points, while the size of the statistical uncertainty is shown by the error bars. The blue arrows represent the data points outside of the frame. The hatched band shows the total systematic uncertainty of the post-fit yield [1].

8.1.1 Pulls of Nuisance Parameters

The resulting NPs values and their uncertainties from fits of a $bb\Phi$ signal with mass $m_\Phi = 480$ GeV are shown in Figure 8.3 when using an Asimov dataset and in Figure 8.4 when using the observed dataset. The results for the corresponding fit to a ggF signal with mass $m_\Phi = 480$ GeV are shown in Figures 8.5 and 8.6 for the Asimov and observed datasets respectively. Only those 15 which have the greatest impact on the Parameter of Interest (PoI) are shown.

By the nature of the construction of the Asimov dataset, described in Section 6.2.1, one expects the fitted value of each NP to be its nominal value, however in the observed dataset these NPs can be pulled up or down. The largest of these pulls is for the `TTBAR_MLLSHAPE` parameter which allows the $t\bar{t}$ $m_{\mu\mu}$ shape to vary. The cause for this is the small number of excess events at high mass in the bTag signal region, where this shape uncertainty is largest. The NP with the second-largest impact on the PoI is the muon reconstruction efficiency systematic, which varies the expected yields of the signal and background. Whilst pulled slightly downwards, this NP is still consistent with its nominal value within the fitted uncertainties. The normalisation factors for the $Z+$jets background have some of the largest impacts on the PoI and one can see that the normalisation of the Z+HF background is pulled upwards from its nominal value by 46% in the observed dataset, as expected from the studies in Section 4.4. The remaining NPs are not significantly pulled from their nominal values.

8.1.2 Local p_0

As discussed in Section 6.1, the p_0 is a measure of the significance of any observed excesses in the experimental data. Figure 8.7a shows the p_0 distribution across the searched mass range for the $bb\Phi$ signal hypothesis and Figure 8.7b shows the same distribution for the ggF signal hypothesis. In each case there are three separate fit configurations: in the combined fit all signal and control regions are included; whereas in the "b-Tag" fit, the SRbVeto region is excluded and vice-versa for the "b-Veto" fit configuration.

The "b-Tag" and "b-Veto" distributions are very similar for the two signal hypotheses, as they have very similar shapes. The combined fit differs, however, in that the $bb\Phi$ signal has a far higher acceptance in the bTag signal region. This can be seen for the ggF signal hypothesis in Figure 8.7b where the p_0 of the combined fit matches the "b-Veto" fit very closely but the $bb\Phi$ combined fit in Figure 8.7a is affected by data in both the b-Tag and b-Veto signal regions.

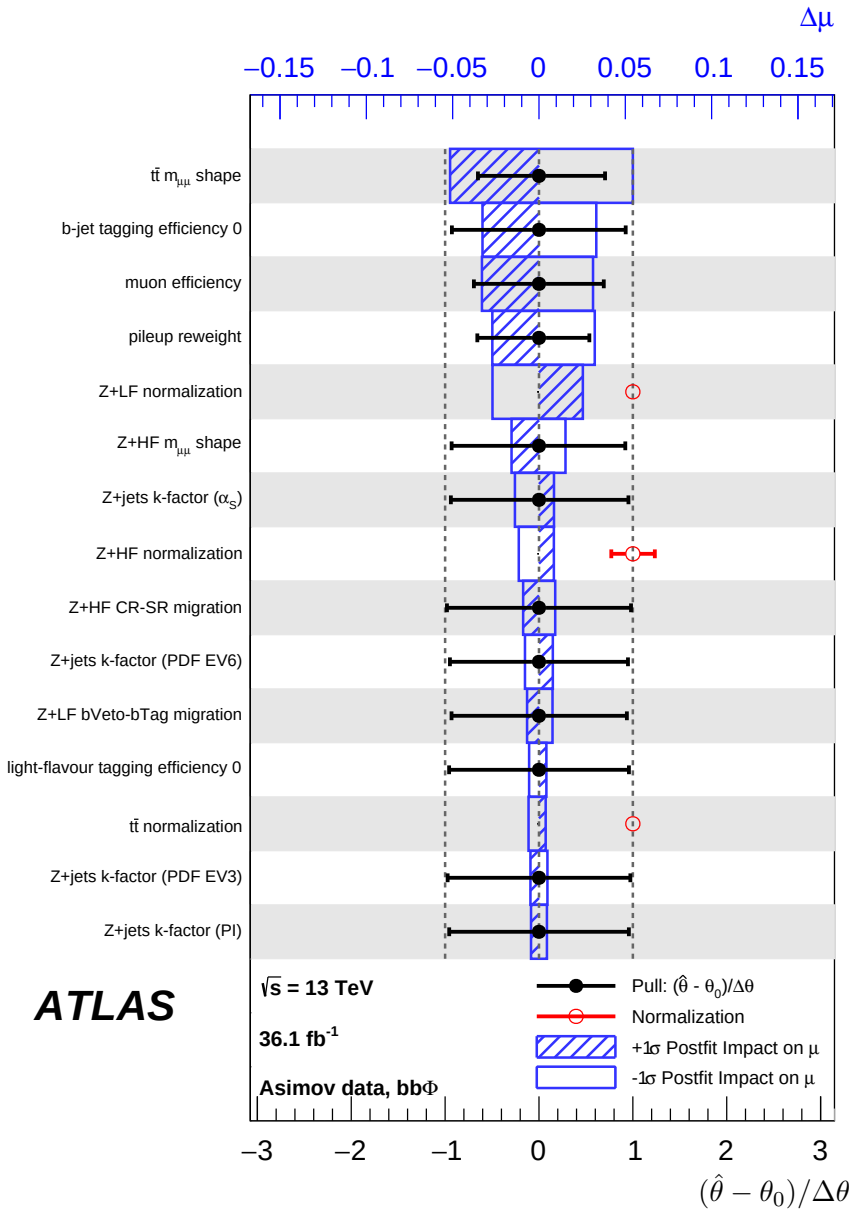


Figure 8.3: Pulls of Nuisance Parameters (NPs) in a fit to a $bb\Phi$ $m_\Phi = 480$ GeV signal hypothesis, using an Asimov dataset. The values of the NPs are shown in black, normalisation factors are shown in red and the impact of these parameters on the PoI is shown in blue. Only the 15 NPs with the largest impact on the PoI are shown.

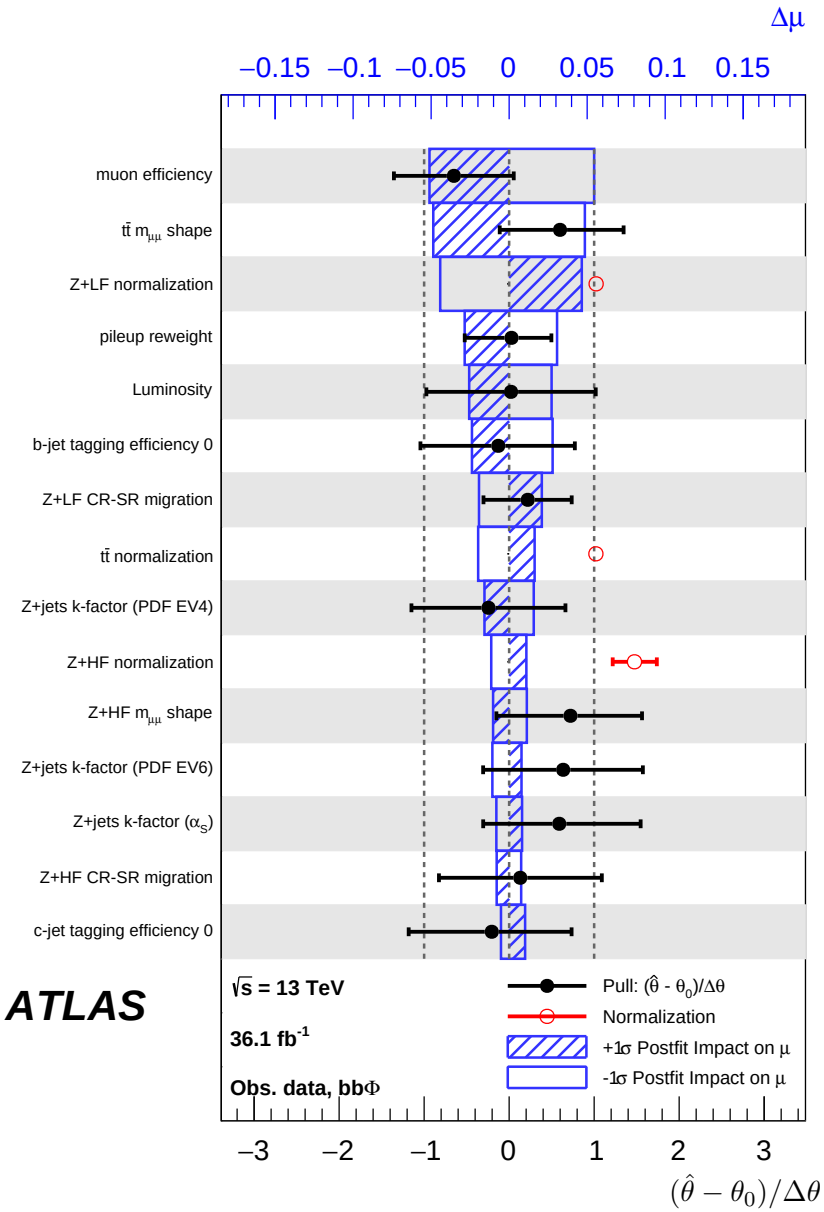


Figure 8.4: Pulls of Nuisance Parameters (NPs) in a fit to a $bb\Phi$ $m_\Phi = 480$ GeV signal hypothesis, using the observed dataset. The values of the NPs are shown in black, normalisation factors are shown in red and the impact of these parameters on the PoI is shown in blue. Only the 15 NPs with the largest impact on the PoI are shown.

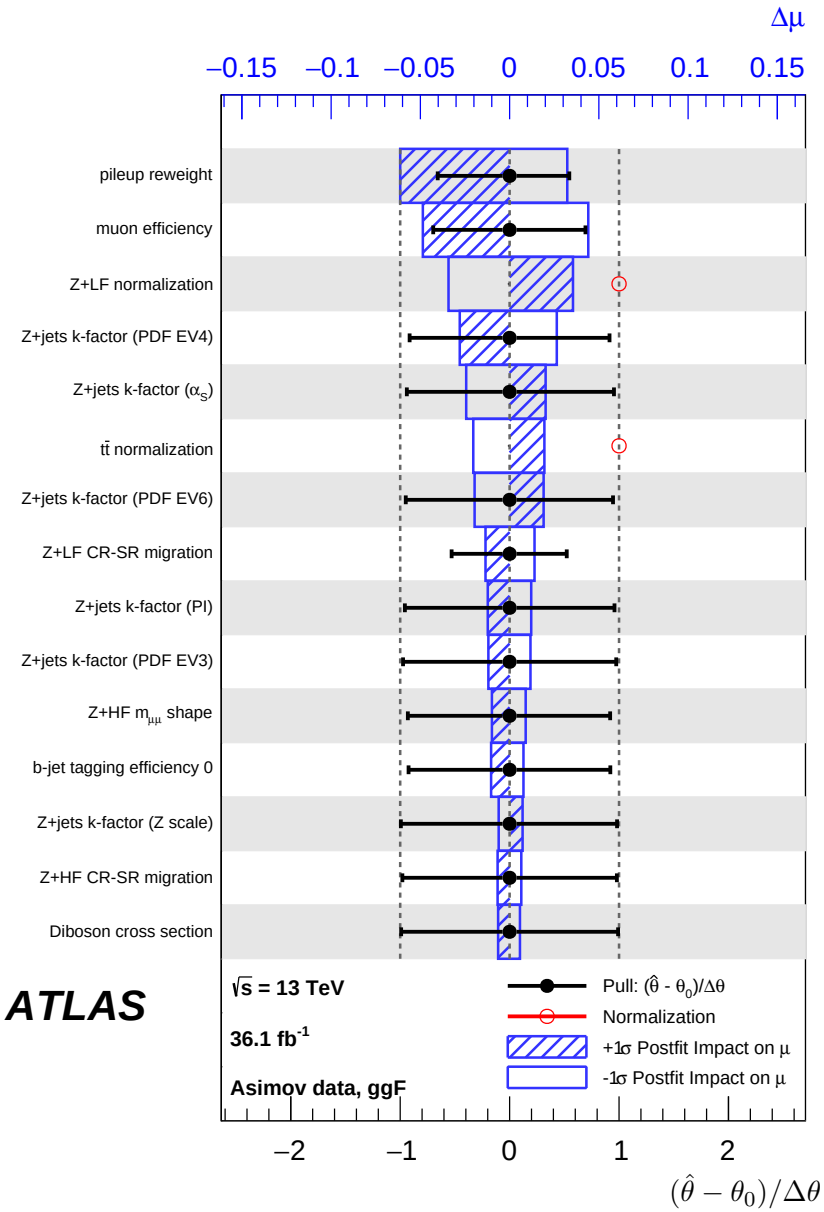


Figure 8.5: Pulls of Nuisance Parameters (NPs) in a fit to a ggF $m_\Phi = 480$ GeV signal hypothesis, using an Asimov dataset. The values of the NPs are shown in black, normalisation factors are shown in red and the impact of these parameters on the PoI is shown in blue. Only the 15 NPs with the largest impact on the PoI are shown.

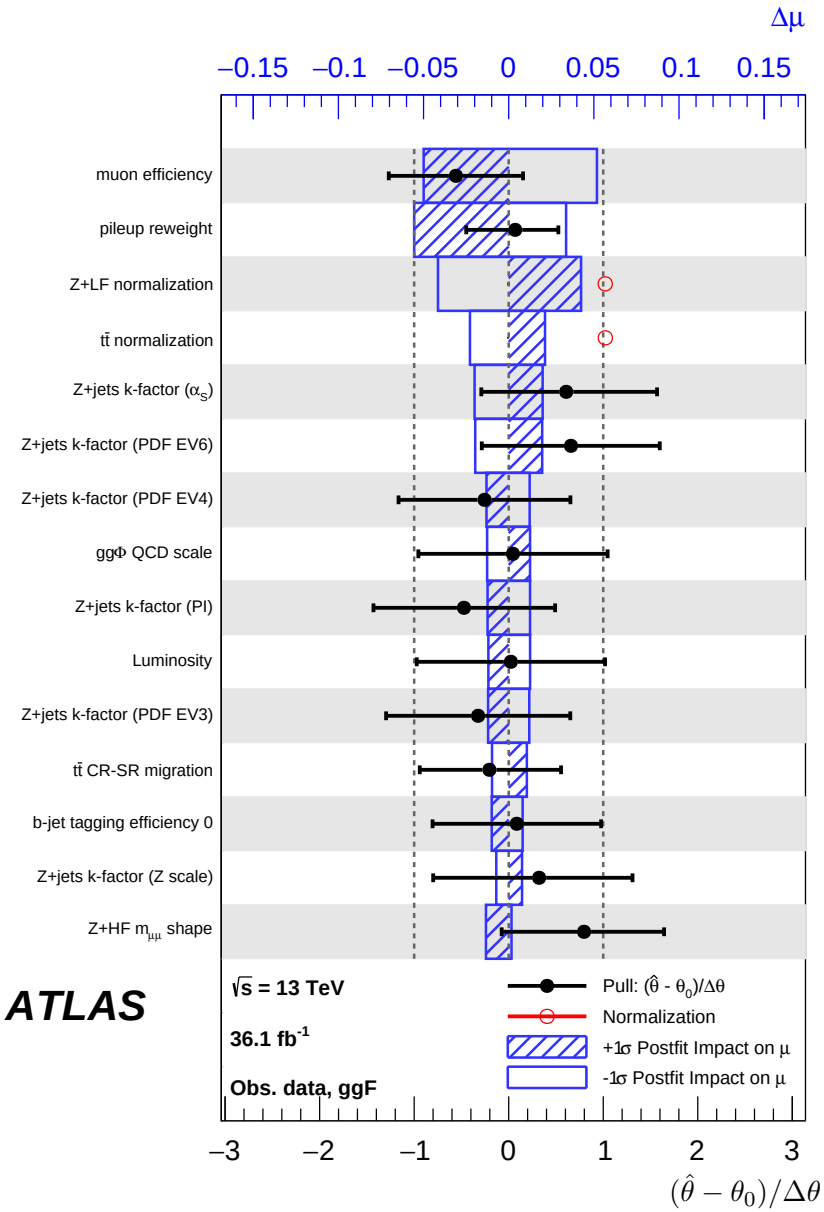
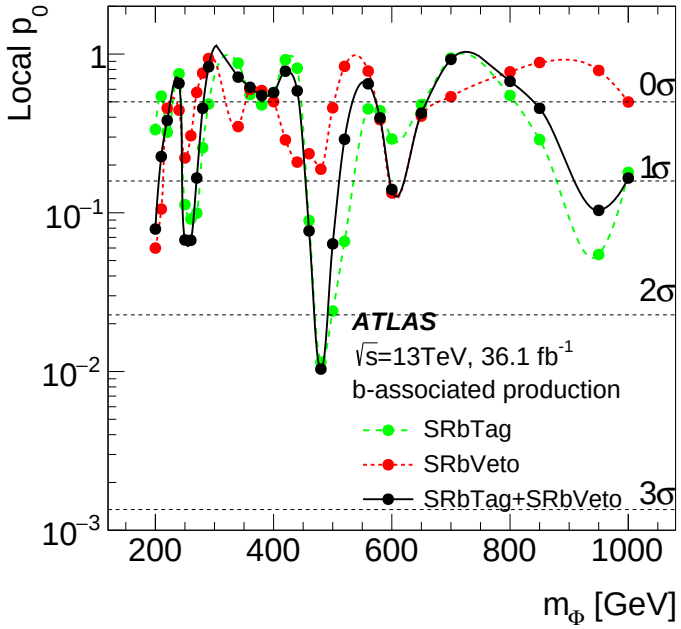
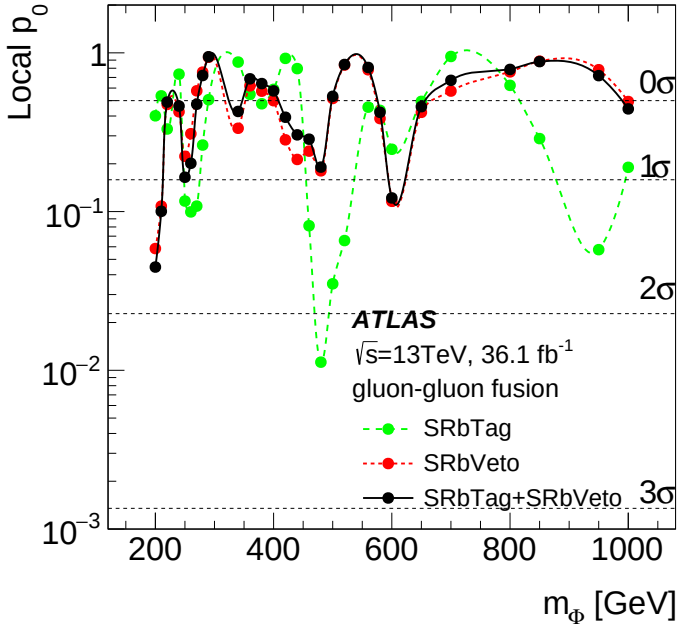


Figure 8.6: Pulls of Nuisance Parameters (NPs) in a fit to a ggF $m_\Phi = 480$ GeV signal hypothesis, using the observed dataset. The values of the NPs are shown in black, normalisation factors are shown in red and the impact of these parameters on the PoI is shown in blue. Only the 15 NPs with the largest impact on the PoI are shown.



(a)



(b)

Figure 8.7: Local p_0 distributions for the (a) $bb\Phi$ and (b) ggF signal hypotheses, with a combined fit (black), and one where only the SRbTag (red) or SRbVeto (green) signal regions have been included in the fit [1].

The most significant excess occurs at approximately 480 GeV with a local significance of 2.37σ for the $bb\Phi$ signal hypothesis. This is predominantly caused by an excess in the bTag region in this mass range. One should note that the excess is not present for the combined fit of the ggF signal, as it has only a very small acceptance to the b-Tag region.

Look Elsewhere Effect

In order to properly quantify the significance of any excesses it is important to recognise that the search covers multiple mass points, increasing the probability of a statistical fluctuation causing a spuriously high significance at one mass point. This is known as the Look-Elsewhere Effect (LEE). A correction for this effect, to find the *global* significance, can be made using [213]:

$$p_{global}(u) = p_{local}(u) + E(u_0)e^{\frac{u_0-u}{2}}, \tag{8.1}$$

where $u = Z^2$, the square of the local significance $Z = 2.37\sigma$ and $E(u_0)$ is the number of times the likelihood ratio test crosses from below to above some significance value, u_0. This is known as the number of *expected upcrossings* [214]. Using $u_0 = 0\sigma$ results in a global significance for the excess at 480 GeV of 0.68σ. Since this is the largest excess observed, there are no statistically significant excesses in this analysis and therefore there is no evidence of new physics.

8.2 Exclusion limits

Upper limits at 95% CL on $\sigma_\Phi \times \mathcal{B}(\Phi \to \mu\mu)$ are calculated as a function of hypothesis mass using the frequentist CL_s method described in Section 6.1 and these are shown in Figures 8.8a and 8.8b. Both the expected and observed limits were calculated using the asymptotic approximation [196] which was validated with MC pseudo-experiments for the $m_\Phi = 1000$ GeV hypothesis. The expected limits for both $bb\Phi$ and ggF production mechanisms are similar, covering the range 25–1.3 fb and 25–1.8 fb respectively for a signal hypothesis mass between 200 GeV and 1 TeV. The observed limits are consistent with the expected limits with the largest deviation occurring at a mass of 480 GeV for the $bb\Phi$ production mechanism.

These exclusion limits have been found with only a minimal set of model assumptions; they are *model independent*. This is an advantage, as these limits can now be easily interpreted and applied to different Beyond the Standard Model (BSM) models. Theorists are able to interpret these results within the context of their models without the need for a full simulation of the ATLAS detector.

One can compare the upper limits for b-associated production in Figure 8.8a with the results of the ATLAS search in the ditau channel, shown in Figure 1.9b. The exclusion limits in the dimuon channel are more stringent - approximately 25 fb compared to 700 fb at $m_\Phi = 200$ GeV and 1.5 fb compared to 4 fb at $m_\Phi = 1$ TeV. This is in large part due to the fact that muons have a far smaller experimental momentum resolution than that of tau leptons.

The higher centre-of-mass energy of the LHC during Run 2 ($\sqrt{s} = 13$ TeV) has also made it possible to exclude signals of greater mass than was previously attainable. The results from this analysis have extended the exclusion limits to $m_\Phi = 1$ TeV, a considerable improvement from analyses of Run 1 data such as that described in [77], where the largest mass considered was $m_\Phi = 500$ GeV.

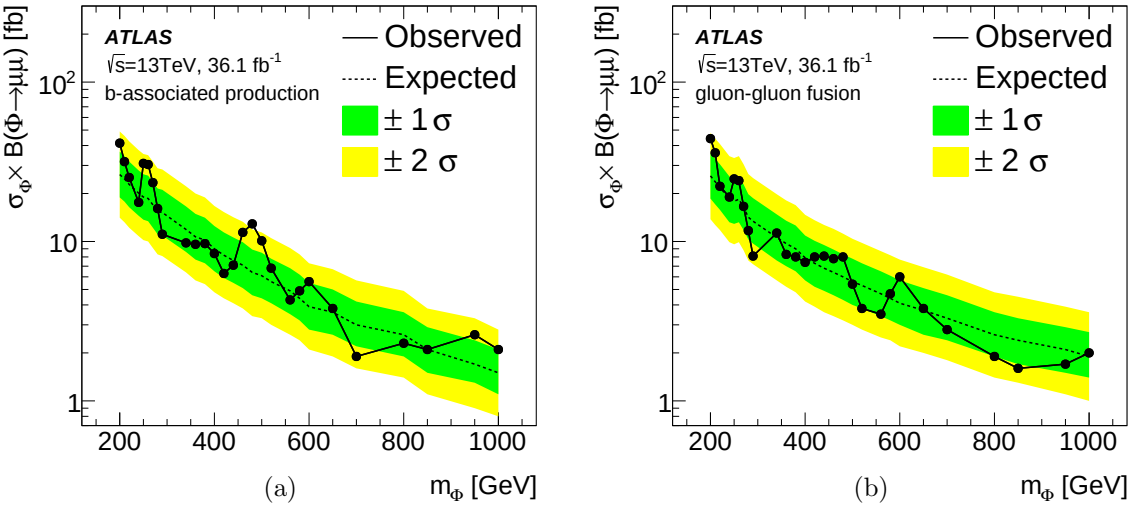


Figure 8.8: Upper limits at 95% CL on $\sigma_\Phi \times \mathcal{B}(\Phi \to \mu\mu)$ for the $bb\Phi$ (a) and ggF (b) signal models. The expected limits are shown with dotted lines and their uncertainties are shown by the coloured bands. The observed limits are shown by the black points [1].

8.3 Conclusion

A search for massive scalar resonances decaying to two opposite-sign muons, in association with a b-jet was performed. The search analysed 36.1 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data that was recorded during 2015 and 2016 with the ATLAS detector at the LHC. The mass range for this search was $0.2 < m_\Phi < 1.0$ TeV and it was assumed that the natural width of the resonance was negligible compared to the experimental resolution. Events were categorised depending on the presence or absence of b-tagged jets to improve signal sensitivity. No significant excesses above

the expected Standard Model background were observed - the largest deviation had a local significance of 2.37σ and a global significance 0.68σ. Upper limits on the cross-section times branching ratio at 95 % confidence level were found for such resonances produced via the gluon-fusion and b-associated production mechanisms. The observed upper limits agreed well with expected results, which range between 25–1.3 fb for b-associated production and 25–1.8 fb for the gluon-gluon fusion production mechanism.

8.4 Recommendations for Future Work

Potential areas of development for this research can be broadly summarised in two categories which could each lead to further sensitivity to new, Beyond Standard Model physics.

8.4.1 Full Run 2 Dataset

LHC Run 2 has recently come to an end, and the total integrated luminosity of the $\sqrt{s} = 13$ TeV dataset now stands at approximately 140 fb^{-1}. The use of the full Run 2 dataset would significantly reduce the statistical uncertainties in the analysis, leading to greater sensitivity, particularly at high masses with smaller event yields. However, there are some challenges that may have to be addressed in order to maximise the benefit of this larger dataset. One salient example is the use of Monte Carlo (MC) simulated event samples. It may soon become prohibitive to generate the necessary quantity of simulated events to match the increasing number of data events. Whilst this could be mitigated somewhat by modifying the procedure that this analysis has used to generate MC samples; in the author's opinion the best course of action is instead to estimate the contributions of Standard Model background processes with functions that can be fitted to data. Although, this would require careful study to ensure the functional form chosen was valid in the entire search region.

8.4.2 Enhanced selection criteria

The brief study of the use of a multivariate selection method, described in Section 5.3, indicated that such a technique could enhance the signal sensitivity by rejecting $t\bar{t}$ background events. These improvements are predicated on the topological differences between the $t\bar{t}$ process and the proxy signal model used in this analysis. There is therefore a trade-off between the analysis sensitivity to a given

model and its applicability to other models. In the future, whilst it would entail a greater workload, it may be advantageous to pursue both *model-dependent* and *model-independent* strategies. This would maximise the constraints on the most "likely" BSM scenarios, whilst remaining open to the possibility that the Universe might have more surprising discoveries in store.

Appendices

A Acronyms

2HDM Two Higgs Doublet Model

4FS 4-Flavour Scheme

BCID Bunch Crossing Identifier

BDT Boosted Decision Tree

BR Branching Ratio

BSM Beyond the Standard Model

CERN Conseil Européen pour la Recherche Nucléaire

CKM Cabibbo-–Kobayashi-–Maskawa

CL Confidence Level

CR Control Region

CTP Central Trigger Processor

CSC Cathode-Strip Chamber

$\boldsymbol{\sigma_{\Phi} \times \mathcal{B}(\Phi \to \mu\mu)}$ Cross-section times Branching Ratio

DIS Deep-Inelastic Scattering

DSCB Double-sided Crystal Ball

EMB Electromagnetic Barrel

EMEC Electromagnetic Endcap

EW Electroweak

EWSB Electroweak Symmetry Breaking

FCal Forward Calorimeter

FSR Final State Radiation

FWHM Full Width at Half Maximum

GSF Gaussian Sum Filter

GUT Grand Unified Theory

HEC Hadronic Endcap Calorimeter

HLT High-Level Trigger

IBL Insertable B-Layer

IRC Infra-red and Collinear safe

ID Inner Detector

ISR Initial State Radiation

JVT Jet-Vertex Tagger

L1 Level 1 - the first stage of the ATLAS trigger

LAr Liquid Argon

LCW Local Hadronic Cell Weighting

LEE Look-Elsewhere Effect

LHC Large Hadron Collider - a large proton collider at CERN near Geneva, Switzerland.

LLR Log-Likelihood Ratio

LO Leading Order

LSP Lightest Stable Particle - the lightest particle in an R-parity conserving SUSY model. Such a particle could be a candidate for Dark Matter.

MC Monte Carlo

MDT Monitored Drift Tube

$\boldsymbol{E_T^{miss}}$ Missing Transverse Energy

MPI Multi-parton Interaction

MS Muon Spectrometer

MSSM Minimal Supersymmetric Model

MVA Multivariate Analysis

NF Normalisation Factor

NLO Next-to-Leading Order

NP Nuisance Parameter

NWA Narrow width approximation

PDF Parton Distribution Function

PI Photon-Induced

PMNS Pontecorvo-Maki-Nakagawa-Sakata

PMT Photomultiplier Tube

PoI Parameter of Interest

PS Proton Synchrotron

PSB Proton Synchrotron Booster

QCD Quantum Chromodynamics

QED Quantum Electrodynamics

QFT Quantum Field Theory

RDO Raw Data Object

RF Radio Frequency

ROC Receiver Operating Characteristics

ROD Read-out Driver

RoI Region of Interest

RPC Resistive Plate Chamber

SCT Semi-conductor tracker

SM Standard Model

SPS Super Proton Synchrotron

SR Signal Region

SUSY Supersymmetry

TES Tau Energy Scale

TGC Thin Gap Chamber

ToT Time over Threshold

TRT Transition Radiation Tracker

TTVA Track-to-Vertex Association

VdM Van der Meer

vev Vacuum expectation value

VR Validation Region

Z*+*HF Z + Heavy Flavour

Z*+*LF Z + Light Flavour

B Systematic Uncertainties

These tables list the systematic uncertainty nuisance parameters. Table B1 lists the muon measurement uncertainties and Tables B3 and B2 list the flavour tagging and jet measurement uncertainties respectively.

Table B1: Summary of muon measurement uncertainties.

Name	Description
MUON_EFF_SYS MUON_EFF_STAT	Uncertainties on the muon reconstruction scale factor
MUON_SCALE	Uncertainty of muon momentum scale
MUON_ID	Variations due to track smearing in the Inner Detector (ID)
MUON_MS	Variations due to track smearing in the Muon Spectrometer (MS)
MUON_TTVA_SYS MUON_TTVA_STAT	Uncertainties on the Track-to-Vertex Association (TTVA) scale factor
MUON_EFF_TrigSystUncertainty MUON_EFF_TrigStatUncertainty	Uncertainties on the trigger efficiency scale factor
MUON_ISO_STAT MUON_ISO_SYS	Uncertainties on the muon isolation scale factor

Table B2: Summary of jet measurement uncertainties.

Name	Description
JET_JER_SINGLE_NP	Uncertainty on Jet Energy Resolution
JET_Flavor_Response	Uncertainty due to differences in gluon jet response between different MC generatrors
JET_Flavor_Composition	Uncertainty due to differences in flavour composition
JET_BJES_Response	Dedicated uncertainty on flavour response for b-jets
JET_EffectiveNP_1 JET_EffectiveNP_2 JET_EffectiveNP_2 JET_EffectiveNP_3 JET_EffectiveNP_4 JET_EffectiveNP_5 JET_EffectiveNP_6 JET_EffectiveNP_7 JET_EffectiveNP_8restTerm	Reduced eigenvector variations from 75 NPs derived from in-situ analyses
JET_Pileup_OffsetMu JET_Pileup_OffsetNPV JET_Pileup_PtTerm JET_Pileup_RhoTopology	Uncertainties on jet energy scale due to pileup effects
JET_EtaIntercalibration_NonClosure JET_EtaIntercalibration_Modelling JET_EtaIntercalibration_TotalStat	Uncertainty on jet η intercalibration procedure
JET_PunchThrough_MC15 JET_PunchThrough_AFII	Uncertainty on modelling of jet punch-through
JET_RelativeNonClosure_AFII	Uncertainty due to use of fast simulation framework
JET_SingleParticle_HighPt	Uncertainty on high p_{T} jets stemming from propagation of single hadron uncertainties to jet

Table B3: Summary of flavour-tagging uncertainties.

Name	Description
FT_EFF_Eigen_B_0	Eigenvector variation for efficiency scale factor for b-quarks
FT_EFF_Eigen_B_1	
FT_EFF_Eigen_B_2	
FT_EFF_Eigen_C_0	Eigenvector variation for efficiency scale factor for c-quarks
FT_EFF_Eigen_C_1	
FT_EFF_Eigen_C_2	
FT_EFF_Eigen_C_3	
FT_EFF_Eigen_Light_0	Eigenvector variation for efficiency scale factor for light quarks
FT_EFF_Eigen_Light_1	
FT_EFF_Eigen_Light_2	
FT_EFF_Eigen_Light_3	
FT_EFF_Eigen_Light_4	
FT_EFF_extrapolation	Extrapolated uncertainty for jets with $p_{\mathrm{T}} >$ 300 GeV
FT_EFF_extrapolation_from_charm	Uncertainty on extrapolation from c-jets in exclusive Soft Muon Tagged (SMT) events to inclusive c-jets

C Expected Yields

Table C1: Expected event yields of $Z+$jets background processes.

$\sqrt{s} = 13$ TeV, $\mathcal{L} = 36$ fb^{-1}	**Zjets**		**acceptance (%)**	
Channel Selection	26170000 ± 6000		100 ± 0	
GRL	26170000 ± 6000		100 ± 0	
SliceMatch	25830000 ± 6000		99 ± 0	
Jet Cleaning	25800000 ± 6000		100 ± 0	
	Light Flavour		**Heavy Flavour**	
Min $M_{\mu\mu} > 60GeV$	24220000 ± 5000	100 ± 0	1331000 ± 1000	100 ± 0
Opposite charge muons	24220000 ± 5000	100 ± 0	1331000 ± 1000	100 ± 0
Trigger Selection	22710000 ± 5000	94 ± 0	1240000 ± 1000	93 ± 0
Trigger Matching	22130000 ± 5000	97 ± 0	1220000 ± 1000	98 ± 0
Isolation Cut	21880000 ± 5000	99 ± 0	1210000 ± 1000	99 ± 0
$p_t^{lead} > 30$ GeV	20970000 ± 5000	96 ± 0	1180000 ± 1000	98 ± 0
$p_t^{sublead} > 30$ GeV	16120000 ± 4000	77 ± 0	728000 ± 1000	62 ± 0
High pT Selection	10720000 ± 4000	67 ± 0	477000 ± 800	66 ± 0
Bad Muon Veto	10660000 ± 4000	99 ± 0	473000 ± 800	99 ± 0
CR Mass Range 100-160 GeV	502000 ± 800	5 ± 0	26200 ± 200	6 ± 0
CRbTag	5900 ± 100	1 ± 0	8900 ± 100	34 ± 0
CRttbar	8 ± 4	0 ± 0	10 ± 5	0 ± 0
CRbVeto	496000 ± 800	99 ± 0	17300 ± 200	66 ± 0
SR Mass Range >160 GeV	58200 ± 200	1 ± 0	3600 ± 60	1 ± 0
SRbTag	1000 ± 40	2 ± 0	1300 ± 30	35 ± 1
SRbVeto	57200 ± 200	98 ± 0	2300 ± 50	65 ± 0

Table C2: Expected event yields of $t\bar{t}$ background processes.

$\sqrt{s} = 13$ TeV, $\mathcal{L} = 36$ fb^{-1},	$t\bar{t}$	acceptance %
Channel Selection	233000 ± 300	100 ± 0
GRL	233000 ± 300	100 ± 0
Zjets slice Match	233000 ± 300	100 ± 0
Jet Cleaning	233000 ± 300	100 ± 0
$M_{\ell\ell} > 60$ GeV	158000 ± 300	68 ± 0
Opposite charge muons	154000 ± 300	98 ± 0
Trigger Selection	144000 ± 200	93 ± 0
Trigger Matching	142000 ± 200	99 ± 0
Isolation Cut	140000 ± 200	98 ± 0
$p_t^{lead} > 30$ GeV	138000 ± 200	99 ± 0
$p_t^{sublead} > 30$ GeV	91100 ± 200	66 ± 0
High pT Selection	59300 ± 200	65 ± 0
Bad Muon Veto	57900 ± 200	98 ± 0
CR Mass Range 100-160 GeV	22600 ± 100	39 ± 0
CRbTag	16200 ± 80	72 ± 0
CRttbar	4100 ± 40	18 ± 0
CRbVeto	2400 ± 30	10 ± 0
SR Mass Range > 160 GeV	18700 ± 90	32 ± 0
SRbTag	16500 ± 80	88 ± 0
SRbVeto	2200 ± 30	12 ± 0

Table C3: Expected event yields of Diboson and Single Top background processes.

$\sqrt{s} = 13$ TeV, $\mathcal{L} = 36$ fb^{-1},	Diboson	acceptance (%)	Single Top	acceptance (%)
Channel Selection	111000 ± 300	100 ± 0	22100 ± 60	100 ± 0
GRL	111000 ± 300	100 ± 0	22100 ± 60	100 ± 0
Zjets slice Match	111000 ± 300	100 ± 0	22100 ± 60	100 ± 0
Jet Cleaning	111000 ± 300	100 ± 0	22100 ± 60	100 ± 0
$M_{\ell\ell} > 60$ GeV	97900 ± 300	88 ± 0	15400 ± 50	70 ± 0
Opposite charge muons	95900 ± 300	98 ± 0	15000 ± 50	97 ± 0
Trigger Selection	88900 ± 300	93 ± 0	14000 ± 50	93 ± 0
Trigger Matching	86600 ± 300	97 ± 0	13900 ± 50	99 ± 0
Isolation Cut	84700 ± 300	98 ± 0	13600 ± 50	98 ± 0
$p_t^{lead} > 30$ GeV	81900 ± 300	97 ± 0	13500 ± 50	99 ± 0
$p_t^{sublead} > 30$ GeV	51800 ± 200	63 ± 0	8990 ± 40	67 ± 0
High pT Selection	34300 ± 200	66 ± 0	5860 ± 30	65 ± 0
Bad Muon Veto	33800 ± 200	99 ± 0	5720 ± 30	98 ± 0
CR Mass Range 100-160 GeV	4900 ± 50	15 ± 0	2190 ± 20	38 ± 0
CRbTag	280 ± 20	6 ± 0	1400 ± 20	62 ± 0
CRttbar	20 ± 2	0 ± 0	300 ± 7	14 ± 0
CRbVeto	4600 ± 50	94 ± 0	540 ± 9	25 ± 0
SR Mass Range > 160 GeV	2790 ± 30	8 ± 0	1980 ± 20	35 ± 0
SRbTag	170 ± 6	6 ± 0	1500 ± 20	75 ± 0
SRbVeto	2620 ± 20	94 ± 0	490 ± 9	25 ± 0

Table C4: Expected event yields of $bb\Phi$ and ggF signal processes.

$\sqrt{s} = 13$ TeV, $\mathcal{L} = 36$ fb^{-1},	$bb\Phi$ 400 GeV	acceptance (%)	ggF 400 GeV	acceptance (%)
Channel Selection	120.46 ± 0.51	100.0 ± 0.0	0.17 ± 0.0	100.0 ± 0.0
GRL	120.46 ± 0.51	100.0 ± 0.0	0.17 ± 0.0	100.0 ± 0.0
Zjets slice Match	120.46 ± 0.51	100.0 ± 0.0	0.17 ± 0.0	100.0 ± 0.0
Jet Cleaning	119.11 ± 0.51	98.89 ± 0.0	0.17 ± 0.0	99.24 ± 0.0
$M_{\ell\ell} > 60$ GeV	119.07 ± 0.51	99.96 ± 0.0	0.17 ± 0.0	99.99 ± 0.0
Opposite charge muons	119.02 ± 0.51	99.96 ± 0.0	0.17 ± 0.0	99.99 ± 0.0
Trigger Selection	116.55 ± 0.5	97.93 ± 0.01	0.16 ± 0.0	98.23 ± 0.01
Trigger Matching	116.47 ± 0.5	99.93 ± 0.0	0.16 ± 0.0	99.99 ± 0.0
Isolation Cut	115.46 ± 0.5	99.13 ± 0.0	0.16 ± 0.0	99.27 ± 0.0
$p_t^{lead} > 30$ GeV	115.45 ± 0.5	100.0 ± 0.0	0.16 ± 0.0	100.0 ± 0.0
$p_t^{sublead} > 30$ GeV	114.74 ± 0.5	99.38 ± 0.0	0.16 ± 0.0	99.45 ± 0.0
High pT Selection	73.16 ± 0.4	63.76 ± 0.13	0.1 ± 0.0	64.17 ± 0.16
Bad Muon Veto	67.98 ± 0.38	92.92 ± 0.04	0.1 ± 0.0	92.31 ± 0.05
CR Mass Range 100-160 GeV	0.09 ± 0.01	0.13 ± 0.02	0.0 ± 0.0	0.11 ± 0.02
CRbTag	0.05 ± 0.01	59 ± 4	0.0 ± 0.0	6 ± 4
CRttbar	-0.0 ± 0.0	0.0 ± 1.63	0.0 ± 0.0	0.19 ± 0.19
CRbVeto	0.04 ± 0.01	40 ± 6	0.0 ± 0.0	93.71 ± 0.9
SR Mass Range > 160 GeV	67.89 ± 0.38	99.86 ± 0.0	0.1 ± 0.0	99.87 ± 0.0
SRbTag	30.28 ± 0.25	44.61 ± 0.21	0.0 ± 0.0	4.62 ± 0.17
SRbVeto	37.6 ± 0.29	55.39 ± 0.19	0.09 ± 0.0	95.38 ± 0.03

Bibliography

[1] ATLAS Collaboration. Search for scalar resonances decaying into $\mu^+\mu^-$ in events with and without b-tagged jets produced in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP* 07 (2019) 117. https://doi.org/10.1007/JHEP07(2019)117.

[2] I Antcheva, M Ballintijn, B Bellenot, et al. ROOT — A C++ framework for petabyte data storage, statistical analysis and visualization. *Comput. Phys. Commun.* 180 (2009) 2499. http://www.sciencedirect.com/science/article/pii/S0010465509002550.

[3] Wouter Verkerke and David P Kirkby. The RooFit toolkit for data modeling. *eConf.* (2003) C0303241. https://arxiv.org/abs/physics/0306116.

[4] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B.* 716 (2012) 1. http://dx.doi.org/10.1016/j.physletb.2012.08.020.

[5] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B.* 716 (2012) 30. http://dx.doi.org/10.1016/j.physletb.2012.08.020.

[6] Walther Gerlach and Otto Stern. Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld. *Zeitschrift für Physik.* 9 (1922) 349. https://doi.org/10.1007/BF01326983.

[7] W Pauli. The Connection Between Spin and Statistics. *Phys. Rev.* 58 (1940) 716. https://link.aps.org/doi/10.1103/PhysRev.58.716.

[8] M. Tanabashi, K. Hagiwara, K. Hikasa, et al. Review of Particle Physics. *Phys. Rev. D.* 98 (2018) 030001. https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[9] J J Thomson M A F.R.S. XL. Cathode Rays. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*. 44 (1897) 293. https://doi.org/10.1080/14786449708621070.

[10] Seth H Neddermeyer and Carl D Anderson. Note on the Nature of Cosmic-Ray Particles. *Phys. Rev.* 51 (1937) 884. https://link.aps.org/doi/10.1103/PhysRev.51.884.

[11] M L Perl, G S Abrams, A M Boyarski, et al. Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation. *Phys. Rev. Lett.* 35 (1975) 1489. https://link.aps.org/doi/10.1103/PhysRevLett.35.1489.

[12] G Danby, J-M. Gaillard, K Goulianos, et al. Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos. *Phys. Rev. Lett.* 9 (1962) 36. https://link.aps.org/doi/10.1103/PhysRevLett.9.36.

[13] K Kodama, N Ushida, C Andreopoulos, et al. Observation of tau neutrino interactions. *Phys. Lett. B.* 504 (2001) 218. http://www.sciencedirect.com/science/article/pii/S0370269301003070.

[14] Belle Collaboration. Test of lepton flavor universality in $B \to K^* l^+ l^-$ decays at Belle. 2019. http://arxiv.org/abs/1904.02440.

[15] Belle Collaboration. Lepton-Flavor-Dependent Angular Analysis of $B \to K^* l^+ l^-$. *Phys. Rev. Lett.* 188 (2017) 111801. https://link.aps.org/doi/10.1103/PhysRevLett.118.111801.

[16] Belle Collaboration. Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method. 2019. http://arxiv.org/abs/1904.08794.

[17] BaBar Collaboration. Measurement of Branching Fractions and Rate Asymmetries in the Rare Decays $B \to K^{(*)} l^+ l^-$. *Phys. Rev. D.* 86 (2012) 032012. http://dx.doi.org/10.1103/PhysRevD.86.032012.

[18] BaBar Collaboration. Evidence for an Excess of $\overline{B} \to D^{(*)} \tau^- \overline{\nu}_\tau$ Decays. *Phys. Rev. Lett.* 109 (2012) 101802. https://link.aps.org/doi/10.1103/PhysRevLett.109.101802.

[19] LHCb Collaboration. Search for lepton-universality violation in $B^+ \to K^+ l^+ l^-$ decays. *Phys. Rev. Lett.* 122 (2019) 191801. http://dx.doi.org/10.1103/PhysRevLett.122.191801.

[20] LHCb Collaboration. Test of lepton universality with $B^0 \to K^{*0} l^+ l^-$ decays. *JHEP* 08 (2017) 055. https://doi.org/10.1007/JHEP08(2017)055.

[21] LHCb Collaboration. Measurement of the ratio of the $B^0 \to D^{*-}\tau^+\nu_\tau$ and $B^0 \to D^{*-}\mu^+\nu_\mu$ branching fractions using three-prong τ-lepton decays. *Phys. Rev. Lett.* 120 (2017) 171802. http://dx.doi.org/10.1103/PhysRevLett.120.171802.

[22] NobelPrize.org. The Nobel Prize in Physics 2015. Accessed 11th March 2019. https://www.nobelprize.org/prizes/physics/2015/summary/.

[23] R. Hagedorn. Remarks on the thermodynamical model of strong interactions. *Nucl. Phys. B.* 24 (1970) 93. http://linkinghub.elsevier.com/retrieve/pii/0550321370900568.

[24] Belle Collaboration. Observation of a Resonancelike Structure in the $\pi^\pm\psi'$ Mass Distribution in Exclusive $B \to K\pi^\pm\psi'$ Decays. *Phys. Rev. Lett.* 100 (2008) 142001. https://link.aps.org/doi/10.1103/PhysRevLett.100.142001.

[25] LHCb Collaboration. Observation of the Resonant Character of the $Z(4430)^-$ State. *Phys. Rev. Lett.* 112 (2014) 222002. https://link.aps.org/doi/10.1103/PhysRevLett.112.222002.

[26] BESIII Collaboration. Observation of a Charged Charmoniumlike Structure in $e^+e^- \to \pi^+\pi^- J/\psi$ at $\sqrt{s} = 4.26$ GeV. *Phys. Rev. Lett.* 110 (2013) 252001. https://link.aps.org/doi/10.1103/PhysRevLett.110.252001.

[27] LHCb Collaboration. Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays. *Phys. Rev. Lett.* 115 (2015) 72001. https://link.aps.org/doi/10.1103/PhysRevLett.115.072001.

[28] M. Gell-Mann. A schematic model of baryons and mesons. *Phys. Lett.* 8 (1964) 214. https://linkinghub.elsevier.com/retrieve/pii/S0031916364920013.

[29] G Zweig. An SU_3 model for strong interaction symmetry and its breaking; Version 1. Technical Report CERN-TH-401, CERN, Geneva, jan 1964. https://cds.cern.ch/record/352337.

[30] G Zweig. An SU_3 model for strong interaction symmetry and its breaking; Version 2. Technical Report CERN-TH-412, CERN, Geneva, feb 1964. https://cds.cern.ch/record/570209.

[31] J. D. Bjorken and E. A. Paschos. Inelastic Electron-Proton and γ-Proton Scattering and the Structure of the Nucleon. *Phys. Rev.* 185 (1969) 1975. https://link.aps.org/doi/10.1103/PhysRev.185.1975.

[32] S. L. Glashow, J. Iliopoulos, and L. Maiani. Weak Interactions with Lepton-Hadron Symmetry. *Phys. Rev. D.* 2 (1970) 1285. https://link.aps.org/doi/10.1103/PhysRevD.2.1285.

[33] J. E. Augustin, A. M. Boyarski, M. Breidenbach, et al. Discovery of a Narrow Resonance in e^+e^- Annihilation. *Phys. Rev. Lett.* 33 (1974) 1406. https://link.aps.org/doi/10.1103/PhysRevLett.33.1406.

[34] J. J. Aubert, U. Becker, P. J. Biggs, et al. Experimental Observation of a Heavy Particle J. *Phys. Rev. Lett.* 33 (1974) 1404. https://link.aps.org/doi/10.1103/PhysRevLett.33.1404.

[35] Makoto Kobayashi and Toshihide Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics.* 49 (1973) 652. https://academic.oup.com/ptp/article-lookup/doi/10.1143/PTP.49.652.

[36] S. W. Herb, D. C. Hom, L. M. Lederman, et al. Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions. *Phys. Rev. Lett.* 39 (1977) 252. https://link.aps.org/doi/10.1103/PhysRevLett.39.252.

[37] F. Abe, H. Akimoto, A. Akopian, et al. Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab. *Phys. Rev. Lett.* 74 (1995) 2626. https://link.aps.org/doi/10.1103/PhysRevLett.74.2626.

[38] S. Abachi, B. Abbott, M. Abolins, et al. Search for High Mass Top Quark Production in $p\overline{p}$ Collisions at $\sqrt{s} = 1.8$ TeV. *Phys. Rev. Lett.* 74 (1995) 2422. https://link.aps.org/doi/10.1103/PhysRevLett.74.2422.

[39] R Brandelik, W Braunschweig, K Gather, et al. Evidence for a spin-1 gluon in three-jet events. *Phys. Lett. B.* 97 (1980) 453. http://www.sciencedirect.com/science/article/pii/0370269380906395.

[40] G Arnison, A Astbury, B Aubert, et al. Experimental observation of lepton pairs of invariant mass around 95 GeV/c^2 at the CERN SPS collider. *Phys. Lett. B.* 126 (1983) 398. http://www.sciencedirect.com/science/article/pii/0370269383901880.

[41] G Arnison, A Astbury, B Aubert, et al. Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}$ =540 GeV. *Phys. Lett. B.* 122 (1983) 103. http://www.sciencedirect.com/science/article/pii/0370269383911772.

[42] P. Duinker. Review of e^+e^- physics at PETRA. *Reviews of Modern Physics.* 54 (1982) 325. https://link.aps.org/doi/10.1103/RevModPhys.54.325.

[43] UA1 Collision Events. 1985. Image scanned from original photo negative on 19 Jun 2015. http://cds.cern.ch/record/1847635.

[44] V Khachatryan, A M Sirunyan, A Tumasyan, et al. Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV. *Phys. Rev. D.* 92 (2015) 12004. https://link.aps.org/doi/10.1103/PhysRevD.92.012004.

[45] G Aad, B Abbott, J Abdallah, et al. Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector. *Eur. Phys. J. C.* 75 (2015) 476. https://doi.org/10.1140/epjc/s10052-015-3685-1.

[46] E Noether. Invariante Variationsprobleme. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse.* (1918) 235. http://eudml.org/doc/59024.

[47] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental Test of Parity Conservation in Beta Decay. *Phys. Rev.* 105 (1957) 1413. https://link.aps.org/doi/10.1103/PhysRev.105.1413.

[48] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π Decay of the K_2^0 Meson. *Phys. Rev. Lett.* 13 (1964) 138. https://link.aps.org/doi/10.1103/PhysRevLett.13.138.

[49] K. Abe, R. Akutsu, A. Ali, et al. Search for CP Violation in Neutrino and Antineutrino Oscillations by the T2K Experiment with 2.2×10^{21} Protons on Target. *Phys. Rev. Lett.* 121 (2018) 171802. https://link.aps.org/doi/10.1103/PhysRevLett.121.171802.

[50] P. A. M. Dirac. The Quantum Theory of the Electron. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences.* 117 (1928) 610. http://rspa.royalsocietypublishing.org/cgi/doi/10.1098/rspa.1928.0023.

[51] Michael E Peskin and Daniel V Schroeder. *An Introduction to quantum field theory*. Addison-Wesley Publishing Company, Reading, USA, 1995.

[52] P. Achard, O. Adriani, M. Aguilar-Benitez, et al. Measurement of the running of the electromagnetic coupling at large momentum-transfer at LEP. *Phys. Lett. B.* 623 (2005) 26. http://linkinghub.elsevier.com/retrieve/pii/S0370269305010439.

[53] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.* 96 (1954) 191. https://doi.org/10.1103/PhysRev.96.191.

[54] H. Fritzsch, M. Gell-Mann, and H. Leutwyler. Advantages of the Color Octet Gluon Picture. *Phys. Lett. B.* 47 (1973) 365. https://doi.org/10.1016/0370-2693(73)90625-4.

[55] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics.* 22 (1961) 579. https://doi.org/10.1016/0029-5582(61)90469-2.

[56] Abdus Salam and J. C. Ward. Weak and electromagnetic interactions. *Il Nuovo Cimento.* 11 (1959) 568. http://link.springer.com/10.1007/BF02726525.

[57] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.* 19 (1967) 1264. https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[58] A. A. Aguilar-Arevalo, B. C. Brown, L. Bugel, et al. Significant Excess of Electronlike Events in the MiniBooNE Short-Baseline Neutrino Experiment. *Phys. Rev. Lett.* 121 (2018) 221801. https://link.aps.org/doi/10.1103/PhysRevLett.121.221801.

[59] T. Aaltonen, V. M. Abazov, B. Abbott, et al. Combination of CDF and D0 W-Boson mass measurements. *Phys. Rev. D.* 88 (2013) 052018. https://link.aps.org/doi/10.1103/PhysRevD.88.052018.

[60] The LEP Collaborations. Measurement of the mass of the Z boson and the energy calibration of LEP. *Phys. Lett. B.* 307 (1993) 187. http://linkinghub.elsevier.com/retrieve/pii/0370269393902109.

[61] Makoto Kobayashi and Toshihide Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics.* 49 (1973) 652. https://dx.doi.org/10.1143/PTP.49.652.

[62] Masami Nakagawa, Shoichi Sakata, and Ziro Maki. Remarks on the Unified Model of Elementary Particles. *Progress of Theoretical Physics.* 28 (1962) 870. https://dx.doi.org/10.1143/PTP.28.870.

[63] A. Einstein. Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik.* 354 (1916) 769. http://doi.wiley.com/10.1002/andp.19163540702.

[64] B. P. Abbott, R. Abbott, T. D. Abbott, et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.* 116 (2016) 061102. https://link.aps.org/doi/10.1103/PhysRevLett.116.061102.

[65] Claus Kiefer. Conceptual Problems in Quantum Gravity and Quantum Cosmology. *ISRN Mathematical Physics.* (2013) 1. https://www.hindawi.com/archive/2013/509316/.

[66] James Bedford. An Introduction to String Theory. jul 2011. http://arxiv.org/abs/1107.3967.

[67] Norbert Bodendorfer. An elementary introduction to loop quantum gravity. jul 2016. http://arxiv.org/abs/1607.05129.

[68] Fritz Zwicky. The Redshift of Extragalactic Nebulae (Die Rotverschiebung von extragalaktischen Nebeln). *Published in Helvetica Physica Acta.* 6 (1933) 110. https://arxiv.org/pdf/1711.01693.pdf.

[69] Vera C. Rubin and Jr. Ford, W. Kent. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *The Astrophysical Journal.* 159 (1970) 379. http://adsabs.harvard.edu/doi/10.1086/150317.

[70] Planck Collaboration, N. Aghanim, Y. Akrami, et al. Planck 2018 results. VI. Cosmological parameters. pages 1–71, 2018. http://arxiv.org/abs/1807.06209.

[71] Antonino Del Popolo. Nonbaryonic Dark Matter in Cosmology. *International Journal of Modern Physics D.* 23 (2014) 1430005. http://www.worldscientific.com/doi/abs/10.1142/S0218271814300055.

[72] A D Sakharov. Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe. *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967) 32. http://dx.doi.org/10.1070/PU1991v034n05ABEH002497.

[73] Jonathan L. Feng. Naturalness and the Status of Supersymmetry. *Annual Review of Nuclear and Particle Science*. 63 (2013) 351. http://www.annualreviews.org/doi/10.1146/annurev-nucl-102010-130447.

[74] P. Fayet. About R-Parity and the Supersymmetric Standard Model. In *The Many Faces of the Superworld*, pages 476–497. WORLD SCIENTIFIC, jul 2000. http://www.worldscientific.com/doi/abs/10.1142/9789812793850_0027.

[75] John Ellis and Keith A. Olive. Supersymmetric Dark Matter Candidates. jan 2010. http://arxiv.org/abs/1001.3651.

[76] M. Carena, S. Heinemeyer, O. Stål, C. E. M. Wagner, and G. Weiglein. MSSM Higgs boson searches at the LHC: benchmark scenarios after the discovery of a Higgs-like particle. *Eur. Phys. J. C.* 73 (2013) 2552. http://link.springer.com/10.1140/epjc/s10052-013-2552-1.

[77] V. Khachatryan, A.M. Sirunyan, A. Tumasyan, et al. Search for neutral MSSM Higgs bosons decaying to $\mu^+\mu^-$ in pp collisions at $\sqrt{s}$ =7 and 8 TeV. *Phys. Lett. B.* 752 (2016) 221. https://doi.org/10.1016/J.PHYSLETB.2015.11.042.

[78] ATLAS Collaboration. Search for additional heavy neutral Higgs and gauge bosons in the ditau final state produced in 36 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP* 01 (2018) 055. https://doi.org/10.1007/JHEP01(2018)055.

[79] LHCb Collaboration. Measurement of the $B_s^0 \to \mu^+\mu^-$ Branching Fraction and Effective Lifetime and Search for $B^0 \to \mu^+\mu^-$ Decays. *Phys. Rev. Lett.* 118 (2017) 191801. http://link.aps.org/doi/10.1103/PhysRevLett.118.191801.

[80] Johannes Haller, Andreas Hoecker, Roman Kogler, et al. Update of the global electroweak fit and constraints on two-Higgs-doublet models. *Eur. Phys. J. C.* 78 (2018) 675. http://dx.doi.org/10.1140/epjc/s10052-018-6131-3.

[81] Heavy Flavor Averaging Group, Y. Amhis, Sw. Banerjee, et al. Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016. *Eur. Phys. J. C.* 77 (2017) 895. http://dx.doi.org/10.1140/epjc/s10052-017-5058-4.

[82] Wolfgang Altmannshofer, Joshua Eby, Stefania Gori, et al. Collider signatures of flavorful Higgs bosons. *Phys. Rev. D.* 94 (2016) 115032. https://link.aps.org/doi/10.1103/PhysRevD.94.115032.

[83] ATLAS Collaboration. Search for high-mass new phenomena in the dilepton final state using proton-proton collisions at $\sqrt{s}$ = TeV with the ATLAS detector. *Phys. Lett. B.* 761 (2016) 372. https://doi.org/10.1016/j.physletb.2016.08.055.

[84] CMS Collaboration. Search for high-mass resonances in dilepton final states in proton-proton collisions at $\sqrt{s}$ = 13TeV. *JHEP* 06 (2018) 120. https://doi.org/10.1007/JHEP06(2018)120.

[85] Howard Georgi and S L Glashow. Unity of All Elementary-Particle Forces. *Phys. Rev. Lett.* 32 (1974) 438. https://link.aps.org/doi/10.1103/PhysRevLett.32.438.

[86] H. Nishino, S. Clark, K. Abe, et al. Search for Proton Decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$. *Phys. Rev. Lett.* 102 (2009) 141801. https://link.aps.org/doi/10.1103/PhysRevLett.102.141801.

[87] Michael E Peskin. Beyond the Standard Model. jun 1997. http://cds.cern.ch/record/326995.

[88] Nima Arkani–Hamed, Savas Dimopoulos, and Gia Dvali. The hierarchy problem and new dimensions at a millimeter. *Physics Letters B.* 429 (1998) 263. http://linkinghub.elsevier.com/retrieve/pii/S0370269398004663.

[89] Oskar Klein. Quantentheorie und fünfdimensionale Relativitätstheorie. *Zeitschrift für Physik.* 37 (1926) 895. http://link.springer.com/10.1007/BF01397481.

[90] Oskar Klein. The Atomicity of Electricity as a Quantum Theory Law. *Nature.* 118 (1926) 516. http://www.nature.com/articles/118516a0.

[91] ATLAS Collaboration. Search for new phenomena in high-mass diphoton final states using 37 fb^{-1} of proton–proton collisions collected at $\sqrt{s}$ =13 TeV with the ATLAS detector. *Phys. Lett. B.* 775 (2017) 105. http://www.sciencedirect.com/science/article/pii/S0370269317308511.

[92] Lisa Randall and Raman Sundrum. Large Mass Hierarchy from a Small Extra Dimension. *Phys. Rev. Lett.* 83 (1999) 3370. https://link.aps.org/doi/10.1103/PhysRevLett.83.3370.

[93] ATLAS Collaboration. Search for the Dimuon Decay of the Higgs Boson in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector. *Phys. Rev. Lett.* 119 (2017) 51802. https://doi.org/10.1103/PhysRevLett.119.051802.

[94] P. Lebrun et al. O.S. Brüning, P. Collier. *LHC Design Report Vol.I: The LHC Main Ring*. Number CERN-2004-003-V-1. 2004. http://cdsweb.cern.ch/record/782076.

[95] P. Lebrun et al. O.S. Brüning, P. Collier. *LHC Design Report Vol.II: The LHC Infrastructure and General Services*. Number CERN-2004-003-V2. 2004. http://cds.cern.ch/record/815187.

[96] M. Benedikt, P. Collier, V. Mertens, J. Poole, and K. Schindl. *LHC Design Report Vol.III: The LHC Injector Chain*. Number CERN-2004-003-V3. 2004. http://cds.cern.ch/record/823808.

[97] J. D. Bjorken. Asymptotic Sum Rules at Infinite Momentum. *Phys. Rev.* 179 (1969) 1547. https://link.aps.org/doi/10.1103/PhysRev.179.1547.

[98] Werner Herr and Bruno Muratori. Concept of luminosity. *CAS - CERN Accelerator School: Intermediate Course on Accelerator Physics*, 2006. https://cds.cern.ch/record/941318.

[99] ATLAS Collaboration. ATLAS Twiki Luminosity Public Results Page. Accessed 11th March 2019. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[100] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC. *Eur. Phys. J. C.* 76 (2016) 653. https://doi.org/10.1140/epjc/s10052-016-4466-1.

[101] Illustration of the LHC Accelerator Complex. Accessed 27th March 2019. https://commons.wikimedia.org/wiki/File:Cern-accelerator-complex.svg.

[102] K Nakamura and Particle Data Group. Review of Particle Physics. *J. Phys G.* 37 (2010) 75021. http://stacks.iop.org/0954-3899/37/i=7A/a=075021.

[103] Christian W. Fabjan and Fabiola Gianotti. Calorimetry for particle physics. *Reviews of Modern Physics.* 75 (2003) 1243. https://link.aps.org/doi/10.1103/RevModPhys.75.1243.

[104] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*. 3 (2008) S08003. https://doi.org/10.1088/1748-0221/3/08/S08003.

[105] M Capeans, G Darbo, K Einsweiler, et al. ATLAS Insertable B-Layer Technical Design Report. Technical report, 2010. http://cds.cern.ch/record/1291633.

[106] Schematic diagram of the ATLAS magnet system. Last Accessed 27th March 2019. http://www.jetgoodson.com/images/thesisImages/magnetSystems.png.

[107] G. Aad, B. Abbott, J. Abdallah, et al. The ATLAS Inner Detector commissioning and calibration. *Eur. Phys. J. C*. 70 (2010) 787. http://link.springer.com/10.1140/epjc/s10052-010-1366-7.

[108] ATLAS Collaboration. *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996. https://cds.cern.ch/record/331061.

[109] L. Aperio Bella. Status of the ATLAS Liquid Argon calorimeter and its performance after two years of LHC operation. *Nucl. Instrum. Meth. A*. 718 (2013) 60. https://linkinghub.elsevier.com/retrieve/pii/S0168900212014180.

[110] ATLAS Collaboration. *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996. https://cds.cern.ch/record/331062.

[111] A. Hrynevich. Performance of the ATLAS Tile Calorimeter. *JINST*. 12 (2017) C06021. https://doi.org/10.1088/1748-0221/12/06/C06021.

[112] M. Aurousseau. The ATLAS Liquid Argon Calorimeter at the LHC: Overview and Performance. *Nuclear Physics B - Proceedings Supplements*. 215 (2011) 110. https://linkinghub.elsevier.com/retrieve/pii/S0920563211002209.

[113] J P Archambault, A Artamonov, M Cadabeschi, et al. Energy calibration of the ATLAS Liquid Argon Forward Calorimeter. *JINST*. 3 (2008) P02002. https://doi.org/10.1088/1748-0221/3/02/P02002.

[114] E Diehl. Calibration and Performance of the ATLAS Muon Spectrometer. Technical Report ATL-MUON-PROC-2011-004, CERN, Geneva, oct 2011. http://cds.cern.ch/record/1385884.

[115] R Hauser. The ATLAS Data Acquisition and High Level Trigger Systems: Experience and Upgrade Plans. Technical Report ATL-DAQ-PROC-2012-073, CERN, Geneva, nov 2012. http://cds.cern.ch/record/1497132.

[116] William Panduro Vazquez. The ATLAS Data Acquisition System in LHC Run 2. *J. Phys. Conf. Series.* 898 (2017) 032017. http://stacks.iop.org/1742-6596/898/i=3/a=032017?key=crossref.77fad68ed90def4434be50c9a16adb7f.

[117] Christoph Eck, J Knobloch, Leslie Robertson, et al. *LHC computing Grid: Technical Design Report.* Technical Design Report LCG. CERN, Geneva, 2005. https://cds.cern.ch/record/840543.

[118] Kyungeon Choi. Tracking and Vertexing with the ATLAS Inner Detector in the LHC Run-2. Technical Report ATL-PHYS-PROC-2017-075, CERN, Geneva, jun 2017. https://cds.cern.ch/record/2271033.

[119] ATLAS Collaboration. Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2. *Eur. Phys. J. C.* 77 (2017) 673. https://doi.org/10.1140/epjc/s10052-017-5225-7.

[120] R. Frühwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth. A.* 262 (1987) 444. https://linkinghub.elsevier.com/retrieve/pii/0168900287908874.

[121] Andreas Salzburger. Optimisation of the ATLAS Track Reconstruction Software for Run-2. *J. Phys. Conf. Series.* 664 (2015) 072042. http://stacks.iop.org/1742-6596/664/i=7/a=072042?key=crossref.c1d153f8aa8896687664b45b57a803d5.

[122] ATLAS Collaboration. Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C.* 76 (2016) 292. https://doi.org/10.1140/epjc/s10052-016-4120-y.

[123] ATLAS Collaboration. ATLAS Twiki Beamspot Public Results. Accessed 11th March 2019. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/BeamSpotPublicResults.

[124] ATLAS Collaboration. Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton-proton collisions at the LHC. *Eur. Phys. J. C.* 77 (2017) 332. https://doi.org/10.1140/epjc/s10052-017-4887-5.

[125] ATLAS Collaboration. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *Eur. Phys. J. C.* 77 (2017) 490. https://doi.org/10.1140/epjc/s10052-017-5004-5.

[126] Gavin P. Salam. Towards jetography. *Eur. Phys. J. C.* 67 (2010) 637. http://www.springerlink.com/index/10.1140/epjc/s10052-010-1314-6.

[127] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP* 04 (2008) 063. https://doi.org/10.1088/1126-6708/2008/04/063.

[128] Ryan Atkin. Review of jet reconstruction algorithms. *J. Phys. Conf. Series.* 645 (2015) 012008. http://stacks.iop.org/1742-6596/645/i=1/a=012008?key=crossref.cf7670068f29afdafa203db94719fd66.

[129] ATLAS Collaboration. Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run. Technical Report ATL-PHYS-PUB-2016-012, CERN, Geneva, jun 2016. https://cds.cern.ch/record/2160731.

[130] ATLAS Collaboration. Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data. Technical Report ATLAS-CONF-2011-102, CERN, Geneva, jul 2011. http://cds.cern.ch/record/1369219.

[131] G Piacquadio and C Weiser. A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS. *J. Phys. Conf. Series.* 119 (2008) 032032. https://doi.org/10.1088/1742-6596/119/3/032032.

[132] ATLAS Collaboration. Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s}$=13 TeV. *JHEP* 08 (2018) 089. https://doi.org/10.1007/JHEP08(2018)089.

[133] ATLAS Collaboration. Flavour tagging data/simulation plots with 2017 software configuration. Accessed 27th March 2019. http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2017-006/.

[134] ATLAS Collaboration. Plots of b-tagging performance before and after the installation of the Insertable B-Layer. Accessed 27th March 2019. http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2017-005/.

[135] ATLAS Collaboration. Performance of the ATLAS muon trigger in 2011. Technical Report ATLAS-CONF-2012-099, CERN, Geneva, jul 2012. http://cds.cern.ch/record/1462601.

[136] ATLAS Collaboration. ATLAS Muon Combined Performance with the full 2016 dataset. Accessed 27th March 2019. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/MUON-2017-001/.

[137] ATLAS Collaboration. Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach. Technical Report ATL-PHYS-PUB-2017-022, CERN, Geneva, dec 2017. https://cds.cern.ch/record/2298955.

[138] R Frühwirth. A Gaussian-mixture approximation of the Bethe–Heitler model of electron energy loss by bremsstrahlung. *Comput. Phys. Commun.* 154 (2003) 131. http://linkinghub.elsevier.com/retrieve/pii/S0010465503002923.

[139] ATLAS Collaboration. Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung. Technical Report ATLAS-CONF-2012-047, CERN, Geneva, may 2012. https://cds.cern.ch/record/1449796.

[140] ATLAS Collaboration. Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton–proton collision data. *Eur. Phys. J. C.* 74 (2014) 2941. https://doi.org/10.1140/epjc/s10052-014-2941-0.

[141] Christian Limbach. Reconstruction and Identification of Tau Leptons in ATLAS. *Nucl. Part. Phys. Proc.* 260 (2015) 195. https://linkinghub.elsevier.com/retrieve/pii/S2405601415000991.

[142] ATLAS Collaboration. Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s}$ = 13TeV. Technical Report ATLAS-CONF-2017-029, CERN, Geneva, may 2017. https://cds.cern.ch/record/2261772.

[143] ATLAS Collaboration. Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC. Technical Report ATL-PHYS-PUB-2015-045, CERN, Geneva, nov 2015. http://cds.cern.ch/record/2064383.

[144] ATLAS Collaboration. Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C.* 78 (2018) 903. https://doi.org/10.1140/epjc/s10052-018-6288-9.

[145] Stefan Weinzierl. Introduction to Monte Carlo methods. jun 2000. http://arxiv.org/abs/hep-ph/0006269.

[146] Stefano Forte, Lluí s Garrido, José I Latorre, and Andrea Piccione. Neural network parametrization of deep-inelastic structure functions. *JHEP* 05 (2002) 062. https://doi.org/10.1088/1126-6708/2002/05/062.

[147] Torbjörn Sjöstrand, Stefan Ask, Jesper R Christiansen, et al. An introduction to PYTHIA 8.2. *Comput. Phys. Commun.* 191 (2015) 159. https://doi.org/10.1016/j.cpc.2015.01.024.

[148] Manuel Bähr, Stefan Gieseke, Martyn A. Gigg, et al. Herwig++ physics and manual. *Eur. Phys. J. C.* 58 (2008) 639. http://www.springerlink.com/index/10.1140/epjc/s10052-008-0798-9.

[149] Bo Andersson, Sandipan Mohanty, and Fredrik Soderberg. Recent Developments in the Lund Model. dec 2002. http://arxiv.org/abs/hep-ph/0212122.

[150] D. Amati and G. Veneziano. Preconfinement as a property of perturbative QCD. *Physics Letters B.* 83 (1979) 87. http://linkinghub.elsevier.com/retrieve/pii/0370269379908967.

[151] Alexander Kupco. Cluster Hadronization in HERWIG 5.9. jun 1999. http://arxiv.org/abs/hep-ph/9906412.

[152] Leif Lönnblad. Correcting the Color-Dipole Cascade Model with Fixed Order Matrix Elements. *JHEP* 05 (2002) 046. https://doi.org/10.1088/1126-6708/2002/05/046.

[153] Nils Lavesson and Leif Lönnblad. W+jets matrix elements and the dipole cascade. *JHEP* 08 (2005) 054. https://doi.org/10.1088/1126-6708/2005/07/054.

[154] S Agostinelli and Others. GEANT4: A simulation toolkit. *Nucl. Instrum. Meth. A.* 506 (2003) 250. https://doi.org/10.1016/S0168-9002(03)01368-8.

[155] J Allison, K Amako, J Apostolakis, et al. Recent developments in Geant4. *Nucl. Instrum. Meth. A.* 835 (2016) 186. http://www.sciencedirect.com/science/article/pii/S0168900216306957.

[156] E. Barberio, J. Boudreau, B. Butler, et al. The GEANT4-Based ATLAS Fast Electromagnetic Shower Simulation. In *Astroparticle, Particle and Space Physics, Detectors and Medical Physics Applications*, pages 802–806. WORLD SCIENTIFIC, jun 2008. http://www.worldscientific.com/doi/abs/10.1142/9789812819093_0133.

[157] Z. Marshall. The ATLAS Simulation Software. *Nuclear Physics B - Proceedings Supplements.* 197 (2009) 254. https://linkinghub.elsevier.com/retrieve/pii/S0920563209008007.

[158] D. Berdine, N. Kauer, and D. Rainwater. Breakdown of the Narrow Width Approximation for New Physics. *Phys. Rev. Lett.* 99 (2007) 111601. https://link.aps.org/doi/10.1103/PhysRevLett.99.111601.

[159] J Alwall, R Frederix, S Frixione, et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP* 07 (2014) 079. https://doi.org/10.1007/JHEP07(2014)079.

[160] M Wiesemann, R Frederix, S Frixione, et al. Higgs production in association with bottom quarks. *JHEP* 02 (2015) 132. https://doi.org/10.1007/JHEP02(2015)132.

[161] Sayipjamal Dulat, Tie-Jiun Hou, Jun Gao, et al. New parton distribution functions from a global analysis of quantum chromodynamics. *Phys. Rev. D.* 93 (2016) 33006. https://link.aps.org/doi/10.1103/PhysRevD.93.033006.

[162] ATLAS Collaboration. ATLAS Run 1 Pythia8 tunes. Technical Report ATL-PHYS-PUB-2014-021, CERN, Geneva, nov 2014. http://cds.cern.ch/record/1966419.

[163] Richard D. Ball, Valerio Bertone, Stefano Carrazza, et al. Parton distributions with LHC data. *Nucl. Phys. B.* 867 (2013) 244. https://doi.org/10.1016/j.nuclphysb.2012.10.003.

[164] D J Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth. A.* 462 (2001) 152. https://doi.org/10.1016/S0168-9002(01)00089-4.

[165] Paolo Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP* 11 (2004) 040. https://doi.org/10.1088/1126-6708/2004/11/040.

[166] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP* 11 (2007) 070. https://doi.org/10.1088/1126-6708/2007/11/070.

[167] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP* 06 (2010) 043. https://doi.org/10.1007/JHEP06(2010)043.

[168] E Bagnaschi, G Degrassi, P Slavich, and A Vicini. Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM. *JHEP* 02 (2012) 088. https://doi.org/10.1007/JHEP02(2012)088.

[169] ATLAS Collaboration. Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP* 09 (2014) 145. https://doi.org/10.1007/JHEP09(2014)145.

[170] Jonathan Pumplin, Daniel Robert Stump, Joey Huston, et al. New Generation of Parton Distributions with Uncertainties from Global QCD Analysis. *JHEP* 07 (2002) 012. https://doi.org/10.1088/1126-6708/2002/07/012.

[171] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. Vector boson plus one jet production in POWHEG. *JHEP* 01 (2011) 095. https://doi.org/10.1007/JHEP01(2011)095.

[172] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. A brief introduction to PYTHIA 8.1. *Comput. Phys. Commun.* 178 (2008) 852. https://doi.org/10.1016/j.cpc.2008.01.036.

[173] N Davidson, T Przedzinski, and Z Was. PHOTOS interface in C++: Technical and physics documentation. *Comput. Phys. Commun.* 199 (2016) 86. https://doi.org/10.1016/j.cpc.2015.09.013.

[174] Charalampos Anastasiou, Lance J Dixon, Kirill Melnikov, and Frank Petriello. High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-leading order. *Phys. Rev. D.* 69 (2004) 94008. https://doi.org/10.1103/PhysRevD.69.094008.

[175] Sergey G Bondarenko and Andrey A Sapronov. NLO EW and QCD proton-proton cross section calculations with mcsanc-v1.01. *Comput. Phys. Commun.* 184 (2013) 2343. https://doi.org/10.1016/j.cpc.2013.05.010.

[176] A D Martin, R G Roberts, W J Stirling, and R S Thorne. Parton distributions incorporating QED contributions. *Eur. Phys. J. C.* 39 (2005) 155. https://doi.org/10.1140/epjc/s2004-02088-7.

[177] Peter Zeiler Skands. Tuning Monte Carlo generators: The Perugia tunes. *Phys. Rev. D.* 82 (2010) 74018. https://doi.org/10.1103/PhysRevD.82.074018.

[178] Michal Czakon and Alexander Mitov. Top++: A program for the calculation of the top-pair cross-section at hadron colliders. *Comput. Phys. Commun.* 185 (2014) 2930. https://doi.org/10.1016/j.cpc.2014.06.021.

[179] M Aliev, H Lacker, U Langenfeld, et al. HATHOR: HAdronic Top and Heavy quarks cross section calculatoR. *Comput. Phys. Commun.* 182 (2011) 1034. https://doi.org/10.1016/j.cpc.2010.12.040.

[180] P Kant, O M Kind, T Kintscher, et al. HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions. *Comput. Phys. Commun.* 191 (2015) 74. https://doi.org/10.1016/j.cpc.2015.02.001.

[181] T Gleisberg, S Höche, F Krauss, et al. Event generation with SHERPA 1.1. *JHEP* 02 (2009) 007. https://doi.org/10.1088/1126-6708/2009/02/007.

[182] Tanju Gleisberg and Stefan Höche. Comix, a new matrix element generator. *JHEP* 12 (2008) 039. https://doi.org/10.1088/1126-6708/2008/12/039.

[183] Fabio Cascioli, Philipp Maierhofer, and Stefano Pozzorini. Scattering Amplitudes with Open Loops. *Phys. Rev. Lett.* 108 (2012) 111601. https://doi.org/10.1103/PhysRevLett.108.111601.

[184] Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert. QCD matrix elements + parton showers: The NLO case. *JHEP* 04 (2013) 027. https://doi.org/10.1007/JHEP04(2013)027.

[185] ATLAS Collaboration. Tagging and suppression of pileup jets with the ATLAS detector. Technical Report ATLAS-CONF-2014-018, CERN, Geneva, may 2014. https://cds.cern.ch/record/1700870.

[186] ATLAS Collaboration. Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of $\sqrt{s} = 7$ TeV. Technical Report ATLAS-CONF-2010-038, CERN, Geneva, jul 2010. http://cds.cern.ch/record/1277678.

[187] A.M. Sirunyan, A. Tumasyan, W. Adam, et al. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *JINST*. 13 (2018) P05011. https://doi.org/10.1088/1748-0221/13/05/p05011.

[188] C. E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*. 27 (1948) 379. http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6773024.

[189] Yoav Freund and Robert E Schapire. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*. 55 (1997) 119. http://linkinghub.elsevier.com/retrieve/pii/S002200009791504X.

[190] Jerome H. Friedman. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*. 29 (2001) 1189. http://projecteuclid.org/euclid.aos/1013203451.

[191] Tianqi Chen and Carlos Guestrin. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*, pages 785–794, New York, New York, USA, 2016. ACM Press. https://doi.org/10.1145/2939672.2939785.

[192] Byron P. Roe, Hai-Jun Yang, Ji Zhu, et al. Boosted decision trees as an alternative to artificial neural networks for particle identification. *Nucl. Instrum. Meth. A*. 543 (2005) 577. https://linkinghub.elsevier.com/retrieve/pii/S0168900205000355.

[193] F Pedregosa, G Varoquaux, A Gramfort, et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 12 (2011) 2825. http://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html.

[194] Pearson's Correlation Coefficient. In *Encyclopedia of Public Health*, pages 1090–1091. Springer Netherlands, Dordrecht. http://www.springerlink.com/index/10.1007/978-1-4020-5614-7_2569.

[195] Philip Sedgwick. Spearman's rank correlation coefficient. In *BMJ (Online)*, volume 349, pages 502–505. Springer New York, New York, NY, 2014. http://www.springerlink.com/index/10.1007/978-0-387-32833-1_379.

[196] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C.* 71 (2011) 1554. https://doi.org/10.1140/epjc/s10052-011-1554-0.

[197] J. Neyman and E. S. Pearson. On the Problem of the Most Efficient Tests of Statistical Hypotheses. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences.* 231 (1933) 289. http://rsta.royalsocietypublishing.org/cgi/doi/10.1098/rsta.1933.0009.

[198] A L Read. Presentation of search results: the CL_s technique. *J. Phys G.* 28 (2002) 2693. https://doi.org/10.1088/0954-3899/28/10/313.

[199] Kyle Cranmer, George Lewis, Lorenzo Moneta, Akira Shibata, and Wouter Verkerke. HistFactory: A tool for creating statistical models for use with RooFit and RooStats. Technical Report CERN-OPEN-2012-016, New York U., New York, jan 2012. https://cds.cern.ch/record/1456844.

[200] Gianpaolo Carlino, Alessandro De Salvo, Andrea Di Simone, et al. ATLAS Muon Calibration Framework. *J. Phys. Conf. Series.* 331 (2011) 072007. https://doi.org/10.1088/1742-6596/331/7/072007.

[201] ATLAS Collaboration. Search for resonances in diphoton events at $\sqrt{s} =$ 13 TeV with the ATLAS detector. *JHEP* 09 (2016) 001. https://doi.org/10.1007/JHEP09(2016)001.

[202] Rene Andrae, Tim Schulze-Hartung, and Peter Melchior. Dos and don'ts of reduced chi-squared. dec 2010. http://arxiv.org/abs/1012.3754.

[203] ATLAS Collaboration. A method for the construction of strongly reduced representations of ATLAS experimental uncertainties and the application thereof to the jet energy scale. Technical Report ATL-PHYS-PUB-2015-014, CERN, Geneva, jul 2015. https://cds.cern.ch/record/2037436.

[204] ATLAS Collaboration. Performance of b-jet identification in the ATLAS experiment. *JINST*. 11 (2016) P04008. https://doi.org/10.1088/1748-0221/11/04/P04008.

[205] ATLAS Collaboration. Performance of algorithms that reconstruct missing transverse momentum in $\sqrt{s} = 8$ TeV proton–proton collisions in the ATLAS detector. *Eur. Phys. J. C*. 77 (2017) 241. https://doi.org/10.1140/epjc/s10052-017-4780-2.

[206] Jun Gao and Pavel Nadolsky. A meta-analysis of parton distribution functions. *JHEP*. 07 (2014) 35.

[207] Jon Butterworth, Stefano Carrazza, Amanda Cooper-Sarkar, et al. PDF4LHC recommendations for LHC Run II. *J. Phys. G*. 43 (2016) 023001. https://doi.org/10.1088/0954-3899/43/2/023001.

[208] Patrick Motylinski, Lucian Harland-Lang, Alan D Martin, and Robert S Thorne. Updates of PDFs for the 2nd LHC run. *Nucl. Part. Phys. Proc.* 273-275 (2016) 2136.

[209] Richard D Ball and Others. Parton distributions for the LHC Run II. *JHEP*. 04 (2015) 40.

[210] ATLAS Collaboration. Simulation of top quark production for the ATLAS experiment at $\sqrt{s}$ =13 TeV. Technical Report ATL-PHYS-PUB-2016-004, CERN, Geneva, jan 2016. https://cds.cern.ch/record/2120417.

[211] ATLAS Collaboration. Comparison of Monte Carlo generator predictions to ATLAS measurements of top pair production at 7 TeV. Technical Report ATL-PHYS-PUB-2015-002, CERN, Geneva, jan 2015. https://cds.cern.ch/record/1981319.

[212] Iain W. Stewart and Frank J. Tackmann. Theory uncertainties for Higgs mass and other searches using jet bins. *Phys. Rev. D*. 85 (2012) 034011. https://link.aps.org/doi/10.1103/PhysRevD.85.034011.

[213] Eilam Gross and Ofer Vitells. Trial factors for the look elsewhere effect in high energy physics. *Eur. Phys. J. C*. 70 (2010) 525. https://doi.org/10.1140/epjc/s10052-010-1470-8.

[214] S. Algeri, D.A. van Dyk, J. Conrad, and B. Anderson. On methods for correcting for the look-elsewhere effect in searches for new physics. *JINST*.

11 (2016) P12010. http://stacks.iop.org/1748-0221/11/i=12/a=P12010?key=crossref.5402d863feb6d99341c809a14f5510d3.